Design and Synthesis of Foldamers Derived from Nucleoside β-Amino Acids

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

By

Richard Neil Threlfall

September 2008

Abstract

Oligomers composed of unnatural building blocks which preferentially adopt defined conformations are known as "foldamers". This project is concerned with the design and synthesis of such foldamers, in the form of β-peptides composed of nucleoside β-amino acids, and the investigation of their structural properties.

Starting from the corresponding natural deoxynucleosides thymidine and 2'-deoxyadenosine, the synthesis of the nucleoside β-amino acid monomers was designed to include easily accessible 3'-azido-3'-deoxy intermediates and also amine and nucleobase protecting groups compatible with solid-phase peptide synthesis (SPPS). β-Peptides were produced from the corresponding nucleoside β-amino acids utilising SPPS protocols optimised to suit these highly unusual amino acids. Homo-oligomers varying in length from 2 to 8 residues were obtained by this method as was a single mixed sequence. Examination of peptides derived from thymidine by TOCSY, NOSEY and VT NMR together with UV melting studies showed the presence of secondary structure in the form of an intramolecular hydrogen bonding network. This was assigned as an 8-helix. The furanose sugars formed a zigzag motif at the core of the helix with the nitrogenous bases located alternately on opposite sides of the scaffold. Unsuccessful deprotection of peptides derived from 2'-deoxyadenosine prevented detailed structural analysis of these oligomers.

As an extension to peptide synthesis studies with nucleoside β-amino acids, initial investigations into the replacement of solid-phase protocols with novel "heavy" fluorous-phase chemistry were undertaken. A large, highly fluorinated dendrimer was used as a "fluorous support" and reactions, such as removal of protecting groups and amino acid coupling were successfully achieved. Products were effectively purified without the use of chromatography and characterised by NMR without cleavage from the fluorous support. These preliminary results soundly demonstrated the proof of principle of nucleoside β-peptide synthesis by fluorous chemistry.

Contents

Acknowledgements

I would like to extend my most profound gratitude to my supervisor Rick Cosstick for all his support, guidance and unwavering enthusiasm throughout the course of this project. I am also extremely grateful to Associate Professor Takeshi Wada and his group at the Graduate School of Frontier Sciences, University of Tokyo for hosting my JSPS summer programme with such eagerness and for all their kind assistance.

I would like to thank Julie Fisher of the University of Leeds and Nicola Haworth of Heriot-Watt University for sharing their expertise in NMR and solid-phase peptide synthesis, as well as all the technical staff at the University of Liverpool chemistry department, especially Allan Mills for mass spectrometry services.

I would also like to acknowledge the Engineering and Physical Science Research Council, the Japan Society for the Promotion of Science and the Daiwa Anglo-Japanese Foundation for their generous financial support.

To all my lab-mates, thanks for making the lab such a fun and friendly place to be.

Finally, I would like to dedicate this work to my family and most of all to my wonderful wife, Junna, for all their love, encouragement and support over the past three years. This is for you.

.

Glossary of Terms

A	Adenine
Ac_2O	Acetic anhydride
ACHC	2-Aminocylohexanecarboxylic acid
ACPC	2-Aminocyclopentanecarboxylic acid
AIBN	Azo*bis*isobutyronitrile
AZT	3'-Azido-3'-deoxythymidine
BAIB	*Bis*-acetoxyiodobenzene
Boc	*tert*-Butoxycarbonyl
BzCl	Benzoyl chloride
C	Cytosine
dA	2'-Deoxyadenosine
DBF	Dibenzofulvene
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCC	Dicyclohexyl carbodiimide
DCM	Dichloromethane
DFQ-COSY	Double quantum filtered correlation spectroscopy
DIAD	Di*iso*propylazodicarboxylate
DIPEA	Di*iso*propylethylamine
DMAc	*N,N,*-Dimethylacetamide
DMF	*N,N,*-Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
ES	Electrospray
Et_2O	Diethyl ether
EtOAc	Ethyl acetate
Fmoc	9-Fluorenylmethoxycarbonyl
FPS	Fluorous-phase synthesis
FPPS	Fluorous-phase peptide synthesis
F-SPE	Fluorous solid-phase extraction
FT-IR	Fourier-transform infra red
G	Guanine
HBTU	*O*-Benzotriazole-*N,N,N',N'*-tetramethyluronium hexafluorophosphate
HIV	Human immunodefficiency virus

HMPB-BHA	4-Hydroxymethyl-3-methoxyphenoxybutyric acid benzhydrylamine
(RP)-HPLC	(Reverse-phase)-High performance liquid chromatography
MALDI-TOF	Matrix assisted laser desorption ionisation-time of flight
mCPBA	*meta*-Chloroperbenzoic acid
Me	Methyl
MeCN	Acetonitrile
MeOH	Methanol
MMFF	Merck Molecular Force Field
mdA	3'-Amino-3'-deoxyadenosyl-5'-carboxylic acid
mT	3'-Amino-3'-deoxythymidyl-5'-carboxylic acid
NaOEt	Sodium methoxide
NaOH	Sodium hydroxide
NMP	*N*-methylpyrrolidone
nOe	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
OMs	Methanesulfonyl
(VT)-NMR	(Variable temperature)-Nuclear magnetic resonance
PMBz	*Para*-methoxybenzoyl
PNA	Peptide nucleic acid
ppbK-1	Parts per billion per Kelvin
ppm	Parts per million
PS	Polystyrene
PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
RNA	Ribonucleic acid
SPPS	Solid phase peptide synthesis
T	Thymine
TEAB	Tetraethylammonium bicarbonate
TEMPO	2,2,6,6-Tetramethylpiperidinyloxy
TES	Triethylsilane
TFA	Trifluoroacetic acid
TfO	Trifluoromethanesulfonic
TLC	Thin layer chromatography
Tm	Melting temperature
TMSCl	Trimethylsilylchloride
TNBS	2,4,6-Trinitrobenzenesulfonic acid

TOCSY	Total correlation spectroscopy
U	Uracil
UV	Ultraviolet

Publications

The following publications prepared during the course of this project are contained within the appendix:

"Foldamers derived from nucleoside β-amino acids: PNA or DNA? Can we have both in one place?" Threlfall, R.; Davies, A.; Howarth, N.; Cosstick, R. *Nucleosides, Nucleotides & Nucleic Acids* **2007**, *26*, 611

"Peptides derived from nucleoside β-amino acids form an unusual 8-helix" Threlfall, R.; Davies, A.; Howarth, N. M.; Fisher, J.; Cosstick, R. *Chem. Comm.* **2008**, 585

"Towards fluorous phase synthesis of nucleoside β-peptides" Threlfall, R.; Cosstick, R.; Wada, T. *Nucleic Acids Symposium Series* **2008**, *52*, 337

"Foldamers derived from nucleoside β-amino acids: a new twist on the DNA helix" Cosstick, R.; Threlfall, R.; Fisher, J. *Nucleic Acids Symposium Series* **2008**, *52*, 313

Chapter 1

Introduction

1 Introduction

A large proportion of natural processes rely on some form of macromolecule with a highly ordered structure that is specially adapted to perform a specific role; the two most obvious examples of this are proteins and nucleic acids exemplified by deoxyribonucleic acid (DNA). Proteins are assemblies of peptide chains which, in turn, are polymers of amino acids formed by condensation of the amine group of one amino acid with the carboxylic acid group of its neighbour creating a "peptide bond". DNA, on the other hand, is an assembly of sugars and nitrogenous bases called nucleosides held together by phosphate bridges between neighbouring units to give polymeric chains, known as polynucleotides. Both peptides and oligonucleotides contain a high concentration of potential H-bond donors and acceptors and consequently a number of intricate structures result from both intra and intermolecular interactions of these hydrogen bonding sites. It is this tendency to adopt distinct conformations and attempts to mimic such properties that has made the study of both of these types of macromolecules a major area of research over the last half a century.

1.1 Primary Structure of DNA

Miescher was the first to physically isolate DNA from white blood cells in 1869 by adding alkali to cells obtained from bandages sourced from his local hospital.[1] He was also the first to postulate that this substance that he christened "nuclein" was a chemical code for biological information, yet curiously he disagreed with Hertwig's statement that nuclein was "responsible for the transmission of hereditary characteristics".[2] In Miescher's view DNA was too simple a molecule to carry genetic information and this had to be the responsibility of more complex protein molecules. By 1900 it was known that DNA consisted of a sugar, phosphate and four nucleobases; adenine, cytosine, guanine and thymine, (A, C, G, and T).

Levene first showed the order of linkage as phosphate to sugar to base in 1929[3], but then went on to suggest that DNA contained equal amounts of each of the four nucleobases and hence something with such little diversity was less likely to contain any chemical code than the richly diverse amino acid sequences in proteins. The advent of X-ray crystallographic techniques in the 1940's culminated in the famous picture of B-form DNA produced by Franklin[4] (see *fig. 1*) which led to Watson and Crick's seminal paper on the double helical structure in 1953.[5]

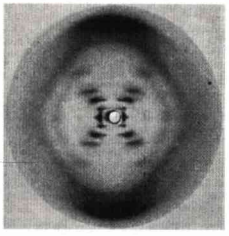


Figure 1: *X-ray diffraction pattern of DNA isolated from calf thymus (taken from reference 4).*

Before considering the conformational properties of a DNA strand it is important to examine the structural features of individual nucleotide units which ultimately define the helix shape adopted. Nucleotides consist of a furanose sugar; a nitrogenous, heterocyclic base and a phosphate ester (see *fig. 2*).

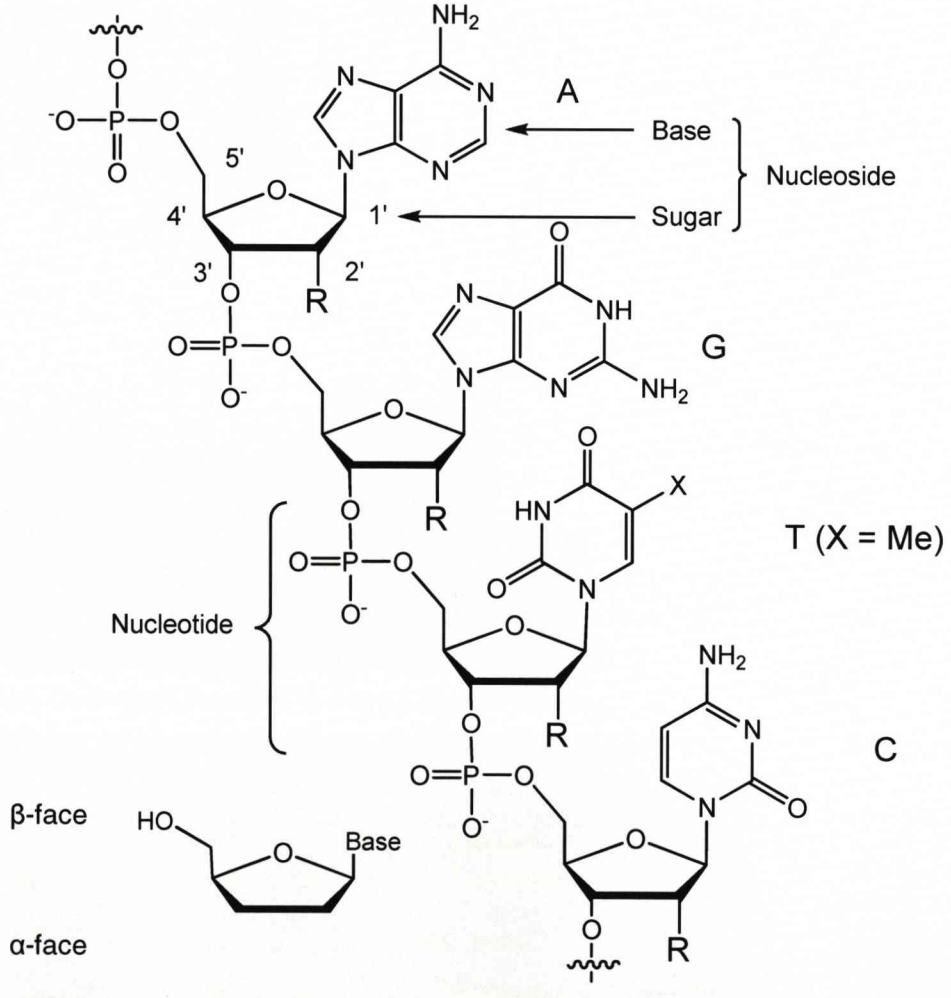


Figure 2: *Chain of 4 nucleotides. In RNA R = OH and T is replaced by U where X = H (figure adapted from reference 6).*

The heterocyclic base is attached to the furanose at the anomeric 1'- position and this sugar/base combination is collectively known as a nucleoside. Heterocyclic bases are bonded on the same face of the sugar ring as the 5'-hydroxymethyl function by the glycosylic bond. This is designated the β-configuration. Bases are either purines (adenine and guanine) or pyrimidines (thymine and cytosine). Pyrimidines are connected to the sugar by the N1 atom, purines by the N9 atom. In ribonucleic acid (RNA), thymine is replaced by uracil; the difference being the lack of the C5 methyl group. Adjacent nucleosides are joined together by the formation of a phosphodiester linkage between the α- configured 3'-oxygen of one nucleoside and the 5'-oxygen of its neighbour.

In single nucleosides the major, helix defining structural features are the position of the base with respect to the sugar, the sugar pucker and the relative positions of C4' and C5' with respect to the rest of the sugar ring. Bases are oriented almost perpendicular to the sugar ring in two main conformations, "*anti*" and "*syn*" (see *fig. 3*), but in the majority of cases nucleosides prefer the *anti* conformation to the *syn.* The *anti* conformation is exclusively adopted by pyrimidines to keep H6 above the ring and avoid the steric clash of O2 with H3'. Adenosine also prefers to be *anti* as this keeps the bulk of the bicyclic base ring away from steric interference with the sugar in favour of the much smaller H8. Guanine, however, has been shown to prefer *syn* conformations in certain circumstances where there is favourable hydrogen bonding between the exocyclic amine and 5'-*O*-phosphate in oligonucleotides.[6] In these cases, *syn* becomes energetically favourable and this results in some unusual structures in G rich oligonucleotides.

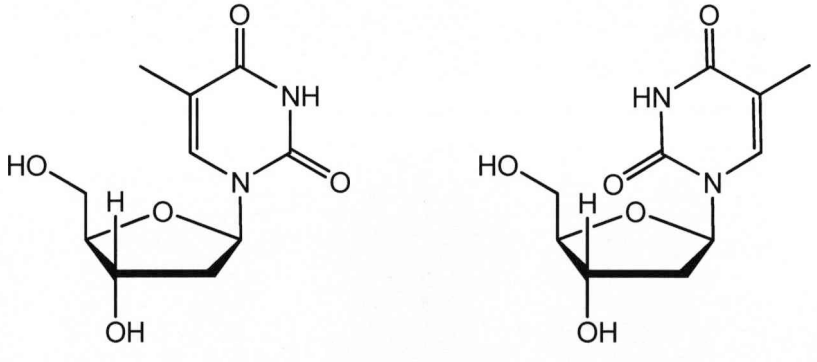


Figure 3: *Favoured anti (left) and disfavoured syn (right) conformations in thymidine.*

Sugar pucker is a description of the displacement of C2' or C3' out of the C1'-O4-C4' plane of the ring and is the major channel by which the non-bonding

interactions of the ring are minimised so as to reduce the overall energy of the sugar. Displacement of either the C2' or C3' carbons towards the same face as the heterocyclic base is referred to as "*endo*", whereas movement towards the lower face is designated "*exo*" (see *fig. 4*). The nomenclature "north" and "south" is used for C3-*endo* and C2'-*endo* respectively as these are the positions of these preferred conformations in the pseudorotation cycle. "N" and "S" are also conveniently the approximate shapes of the C1'-C2'-C3'-C4' bonds (see *fig 4*) and the configurations are shown to be in rapid equilibrium in solution due to the low energy barrier of 20 $kJmol^{-1}$ between them.

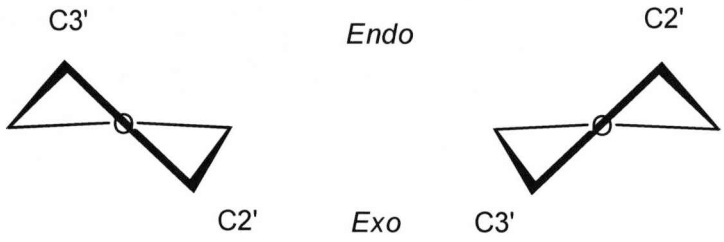


Figure 4: *C3'-Endo ("North", left) and C2'-endo ("South", right) sugar puckers.*

The average or preferred conformation is determined by the orientation of the base and also electronic effects within the ring itself. DNA generally prefers the S-type pucker[7] but whilst up to 95 % of purines are found in this conformation, pyrimidines, especially cytidine, can adopt N-type puckers with a frequency of up to 40 %.[8]

1.2 DNA and RNA Helices

The structure proposed by Watson and Crick for B-form DNA is a right-handed double helix composed of two strands of DNA running anti-parallel to each other and held together by hydrogen bonding between nucleobases on opposite strands. Hydrogen bonds between only specific pairs of bases, adenine and thymine and cytosine and guanine, accounted for the observation of Chargaff that DNA contained approximately equal amounts of the four nucleobases[9] (see *fig. 5*). Adenine and thymine anneal together with two hydrogen bonds and cytosine and guanine have an extra hydrogen bond and are, therefore, more strongly bound than A-T pairs.[10] π- π stacking of the heterocyclic bases within the helix structure is also a major stabilising influence.[10]

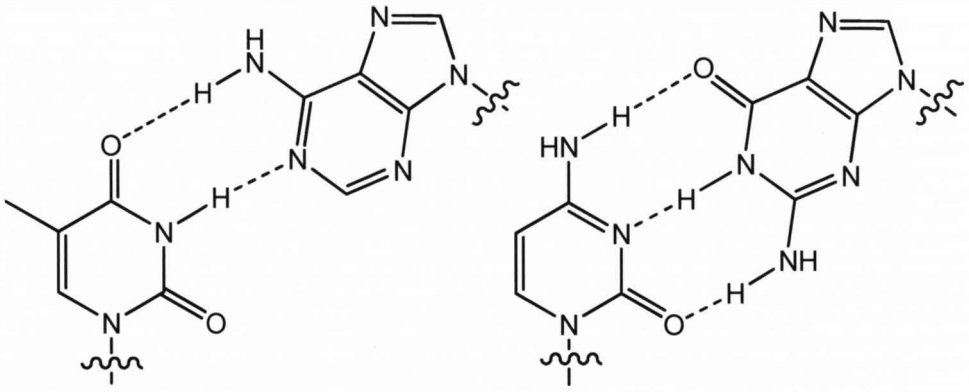


Figure 5: *Watson-Crick hydrogen bonding between A-T (left) and C-G (right).*

Oligonucleotide helices exist in several forms and have structural features common to all forms of helix (see *fig. 6*). There are two "grooves" running down the helical axis at 180 ° to each other, one large, the major groove, and one smaller, the minor groove. The difference arises due to the bonds from sugar to base not being exactly opposite each other (see *fig. 5*)

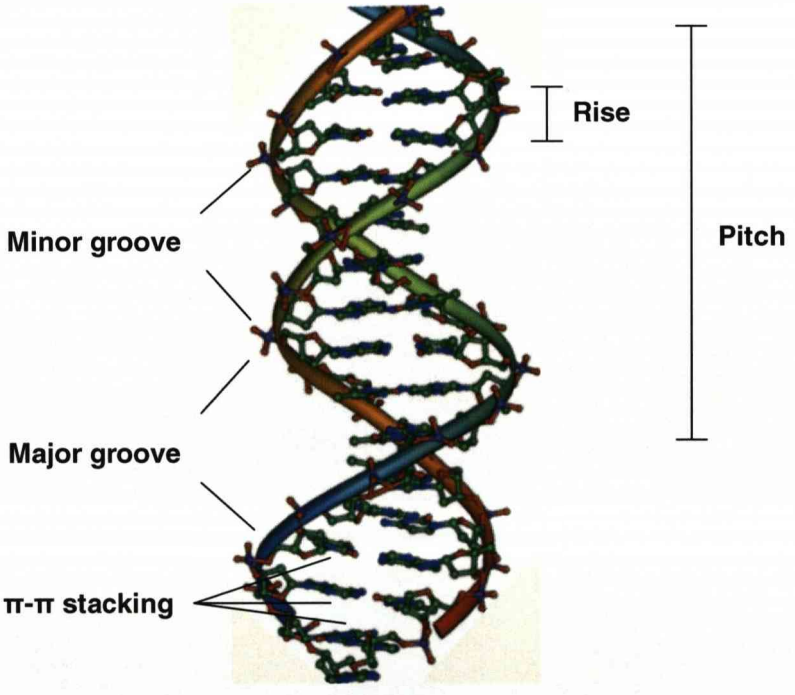


Figure 6: *Representation of a DNA double helix (taken from reference 11).*

There are three main classes of double helix, A, B and Z (see *fig. 7*) and numerous polymorphic substructures which are broadly related have also been observed. The right-handed helix of A-form DNA has anti-parallel strands running from O3' – P – O5' with the bases displaced 4.5 Å from the helical axis. This results in a hollow

core which is 3 Å in diameter. The mean length, or pitch, of an A-type helix measures 24.6 Å and has 11 base pairs per turn with a rotation of 32.7 ° per base pair around the axis. The major groove measures 2.7 Å in width but this slim opening is far deeper than the B-form at 13.5 Å[12] and its sugar pucker is predominantly C3'-*endo*. The slim but deep major groove naturally produces a complimentary wide but shallow minor groove, at 11.0 Å wide but only 2.8 Å deep. A-type helices are also slightly wider than B helices at 26 Å across.

In contrast with the A-form, B-form DNA has a wide major groove, 11.7 Å, and a narrow minor groove, 5.7 Å, both of comparable depth, 8.8 Å and 7.5 Å respectively, and is also a right-handed helix. A full turn of the helix contains 10 base pairs and results in a turn of 36 ° around the helix axis for every base pair.[13] The furanose sugar has a predominantly C2'-*endo* pucker and full turn of a B-form helix has a pitch of 33.2 Å; longer than A-form but shorter than the elongated Z-form, and the diameter of the structure measures 20 Å across.

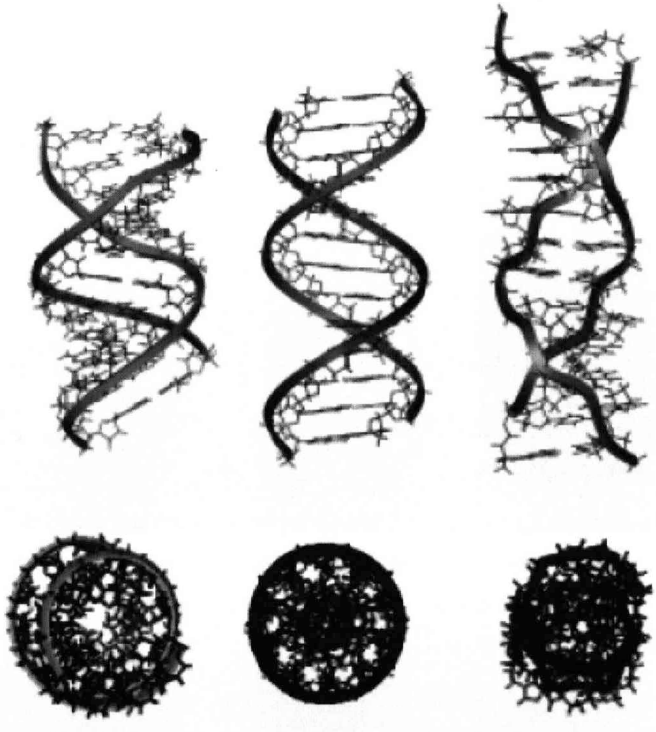


Figure 7: *Comparison of A-form (left), B-form (centre) and Z-form (right) helices (taken from reference 14).*

Z-Form DNA helices have a much more unusual shape than A- or B-form helices and are regarded as a transient features occasionally induced by biological

activity rather than a permanent structure.[15] Although more rare than the other forms of DNA, it was actually the first of the main forms to have its structure solved by x-ray crystallography; by Rich and co-workers in 1979.[16] They found a rather more inelegant structure than the A- or B-form. This time, the helix was left-handed, long and thin, but retained the anti-parallel arrangement of strands. Two nucleotides formed the repeat unit giving the Z-form a zigzag conformation with a pitch of 45.6 Å and 12 base pairs forming a single helix turn. Further studies of single crystals of small d(CGCG) sequences have also found this helix conformation.[17] The conformational switch occurs as a result of a change in glycosylic torsion angle from *anti* to *syn* of one purine of a base pair. For a Z-form helix, every other base must maintain this *syn*-conformation and hence Z-helices are found in regions with alternating purine-pyrimidine sequences, particularly d(C-G)[18] as the pyrimidines preferentially adopt the *anti*-conformation due to steric factors, leaving the purine to switch to *syn*. Sugar pucker is, therefore, correspondingly C3'-*endo* at purine and C2'-*endo* at pyrimidine.

Of the three main forms of DNA discussed, the B-form is the most common seen under normal cellular conditions,[19] followed by A then Z is found in more specialised cases. The A-type helix is less common than its B-DNA counterpart and is found in RNA duplexes,[20] helices containing DNA/RNA hybrids[21] and also in DNA/enzyme complexes.[22] It is not known what the true function of Z-DNA is in the cellular environment, but it is thought that it may be involved in regulation of gene transcription[23] and several Z-DNA binding proteins have also been identified including nuclear-RNA-editing enzyme ADAR1 and the EL3 protein, responsible for the pathogenicity of vaccinia virus.[15]

1.3 Higher Order DNA Structures

Not long after the popularisation of the double helix structure, a number of other structures involving multiple strands of DNA were beginning to emerge. An important feature of structures comprised of multiple strands of DNA is another form of inter-strand hydrogen bonding known as Hoogsteen base pairing.[24] In Watson-Crick base pairs thymine and adenine associate by hydrogen bonding between the thymine N3-H and adenine N3 as well as the thymine C4 carbonyl and adenine N6 amino group. However, in Hoogsteen pairs the thymine N3-H switches hydrogen bonding partner to adenine N7 (see *fig. 8*). This is a major structural feature of many of the more complex DNA structures as it allows unusual modes of binding, for instance, that of an additional single strand to duplex DNA in the major groove.

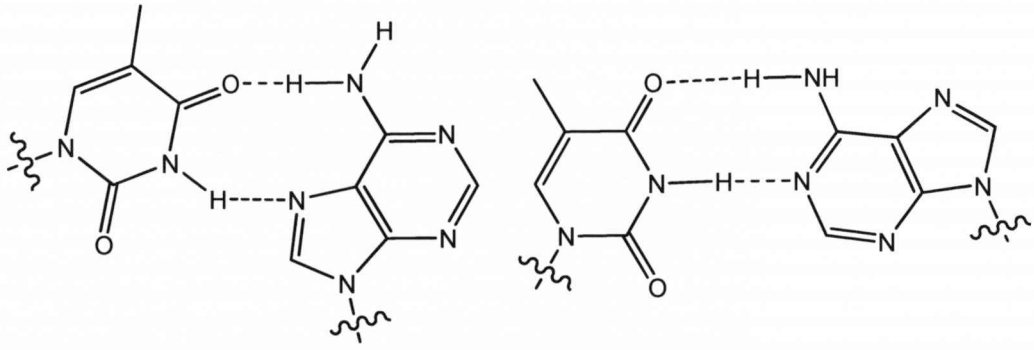


Figure 8: *Comparison of Hoogsteen (left) and Watson-Crick (right) A-T base pairs.*

This arrangement is adopted by the closest related structure to the classical DNA duplex, the triplex. Here, an additional strand of DNA anneals to a duplex in either parallel, in the case of homopyrimidine strands, or anti-parallel, for homopurine strands, fashion.[25] Other, more specialised DNA substructures include the Holliday Junction, G-quadruplex and the *i*-motif. The Holliday junction (see *fig 9.*) was proposed in order to account for the anomaly in genetic recombination that allowed genes on the same chromosome to segregate from each other to produce gametes with differing ratios of the parent gene.[26] The crossing strands can be arranged in either an open planar or anti-parallel conformation[27] and this depends on the presence or absence of Mg^{2+} ions.

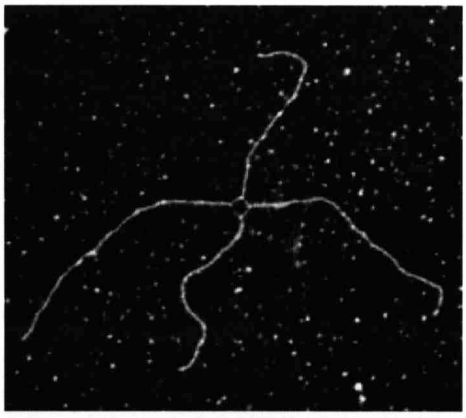
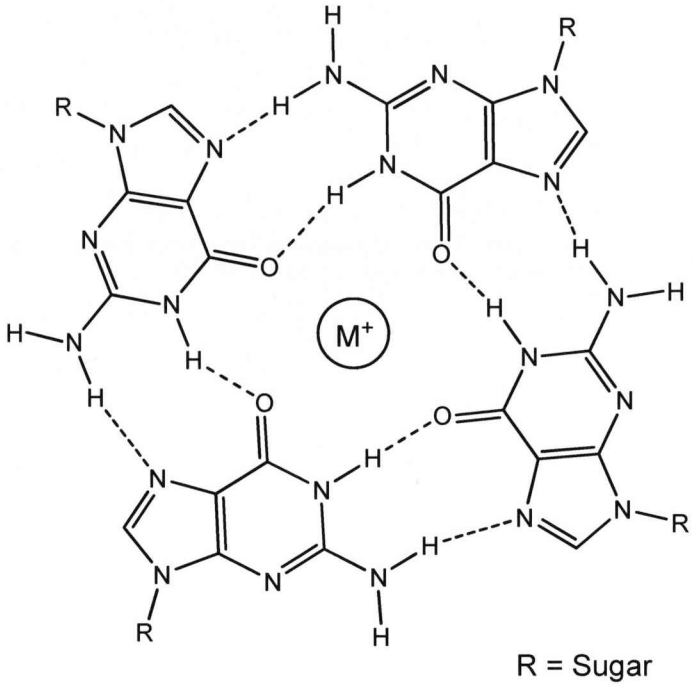


Figure 9: *Electron microscope image of a Holliday junction (taken from reference 29)*[28].

G-quadruplexes, or G-tetraplexes, are formed in G rich sequences of DNA and are assemblies of G-tetrads; square-like formations of four guanine bases stabilised by Hoogsteen hydrogen bonding and also by the presence of monovalent metal ions,

especially potassium,[29] co-ordinating between the plane of stacked tetrads (see *fig. 10*).



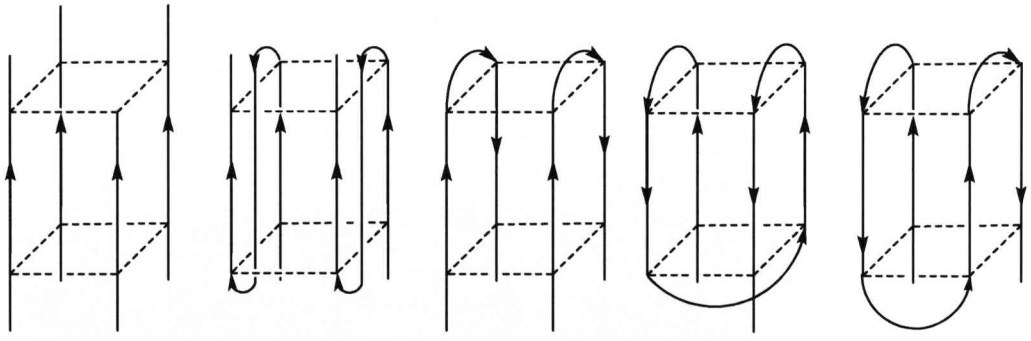


Figure 10: *Structure of a G-tetrad and some folding topologies of G-quadruplexes (adapted from reference 31).*

The G rich sequences can come from four different strands, two strands or even be different sections of the same single strand (see *fig. 10*) but all have different topologies depending on their constitution.[30] G-quadruplexes are found in various places in the human and other genomes and Balasubramanian *et al.* have recently suggested that up to 376,000 sequences within the human genome have the potential to form such structures.[31] Of particular interest are G-quadruplexes which are found at the end of eukaryotic chromosomes, termed telomeres.[32] Here the extreme 3'- end of the telomeric sequence is single stranded, between 100 –

200 nucleotides in length and G-quadruplex formation in human telomeres has been shown to inhibit the activity of the telomerase enzyme[33] which has been implicated in up to 85 % of cancers.[34]

The most unusual higher order DNA structure is the *i*-motif. Formed in cytidine rich strands complementary to the G-rich strands found in telomeric sequences, two parallel duplexes intercalate and the assembly is held together by hemi-protonated C-C^+ base pairs[35] (see *fig 11*).

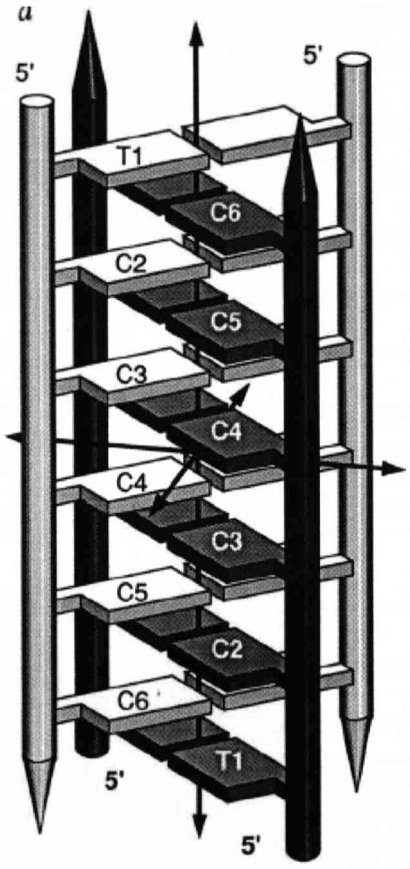


Figure 11: *Representation of an i-motif (taken from reference 36).*

The stability of the i-motif is thus pH dependent; it is most stable at or near the pK_a of cytosine but is less stable at neutral pH.[36] The two interacting duplexes run antiparallel to each other and, as with the G-quadruplex, the four strands of the duplex can be sourced from four single strands,[35] two hairpins[37] or a single strand with multiple C-rich sequences.[38] The biological function of the i-motif is not fully understood, however, several proteins that specifically bind telomeric C-rich sequences have recently been identified.[39]

1.4 Secondary Structure in α-Peptides

The twenty common, genetically coded amino acids observed in nature are known as α-amino acids, that is, the carboxylic acid functionality is attached to the same carbon as the amino moiety (see *fig. 12*). As carboxylic acids and amine groups have a high propensity for hydrogen bonding, when combined with other α-amino acids to give peptides, these peptide chains will fold up to give distinct, complex secondary structures, such as α-helices and β-sheets.

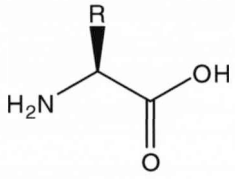


Figure 12: *An α-amino acid.*

Secondary structure in peptides was first proposed by Astbury in 1931 when he observed significant differences in the x-ray fibre diffraction patterns of stretched and non-stretched natural fibres such as wool.[40] He suggested that there were two forms of protein molecules, α- and β-. α- Was the form present in the unstretched fibre and was a coiled helix conformation; β- was produced when stretching the fibre, causing the helix to uncoil. Although inaccurate, his models were extraordinarily close to being correct and formed the basis of the later work by Pauling and co-workers in which the key to secondary structure, the planar peptide bond, was postulated.[41] Astbury's nomenclature was retained and in a later paper to appear in 1933 he also tentatively identified structures in keratins which are today called β-sheets.[42]

It was predicted by Pauling that the six atoms involved in the peptide structure must be in a single plane to support the resonance forms of the whole unit (see *fig. 13*). This, in turn, constrained the conformation of the entire peptide chain which defined limited sites for intramolecular hydrogen bonding to occur. In collaboration with Corey he proposed that in an α-helix the amino acids were arranged with a rise of 5.4 Å with 3.7 residues per turn and hydrogen bond lengths of 2.72 Å.[43] He also proposed that the other major structure that was possible, given the constraints of the planar amide structure, was a helix with 5.1 residues per turn and this became known as the γ-helix[41] which has not been observed to date.

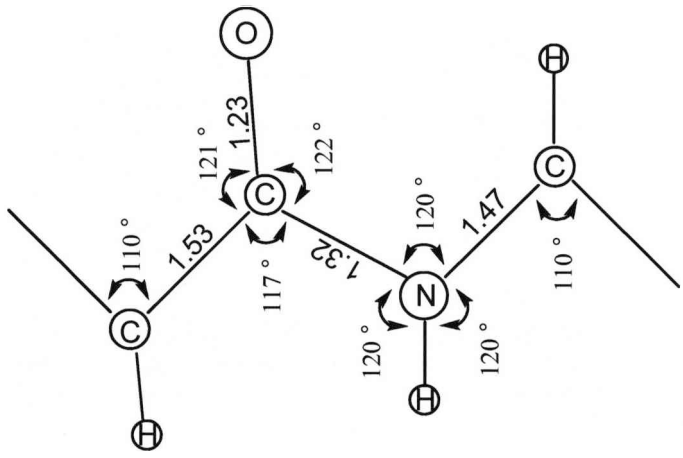


Figure 13: *Pauling's planar amide structure (adapted from reference 42).*

Pauling's α-helix model was, however, not entirely correct. The helix was depicted as left-handed with all the amino acids in the *D*-form, the opposite enantiomer of the natural *L*-form. We now know the reverse of this to be true. Today the accepted dimensions for an α-helix are a pitch of 5.4 Å with 3.6 residues per turn and n, n + 4 hydrogen bonding, that is, the hydrogen bonding occurs between carbonyl groups and amide protons separated by four residues in the direction of the C-terminus (see *fig. 14*). It is thought that around 31 % of all amino acids with secondary structure are in the α-helix form.[44] Variations of α-helices that have been observed include the 3_{10} helix which accounts for approximately 4 % of amino acid structures. This is a tighter helix where there are 3.2 residues per turn and hydrogen bonding is reduced to n, n + 3 forming a 10-membered hydrogen-bonded ring.[45] The π helix, where hydrogen bonding increases to n, n + 5, is considered rare and unstable.[44]

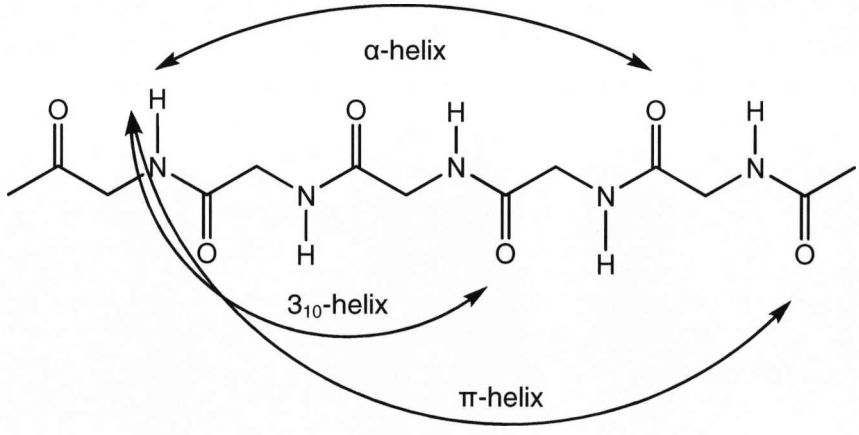


Figure 14: *Hydrogen bonding patterns in α-peptides.*

There is much debate over the basis of the stability of α-helices in aqueous solution and it has been suggested that there are up to 10 major empirical effects that affect this property, including pH dependent formation of salt bridges between side chains, charge stabilisation of the dipole and interaction with metal ions.[46] The role of the solvent is also vitally important when considering the folding propensities of peptide sequences. Before the advent of peptide design it was rare to find short and stable α-helices with aqueous solubility outside of the protein environment,[47] as it is easy for polar solvents such as water to destructively interfere with intramolecular hydrogen bonding. However, even with advances in molecular modelling techniques and production and study of many *de novo* designed peptides and peptidomimetics,[48] the complex relationship between all of the factors involved in α-helix folding is still not fully understood.

The biological role of the lesser observed classes of helices are the subject of some speculation; one theory is that the 3_{10} helix is an intermediate in the folding pathway of a conventional α-helix,[49] but it is the α-helix itself that is the most well understood. Properties of the α-helix can be split into two distinct groups, those derived from the conformational characteristics and those which relate to the electronic features such as dipole. The dipole of an α-helix is derived from the almost linear orientation of the C=O and N-H bonds along the axis of the helix by hydrogen bonding, resulting in alignment of 97 % of the individual amino acid dipoles towards the *N*-terminus.[50] It is thought that this charge differential plays a key role in processes such as ion transport and crystal packing[51] and that as many as 1 in 4 enzymes have a helix dipole influencing the active site.[52] Phosphate binding proteins utilise positively charged side chains at the N-terminus to bind biologically important compounds such as adenosine triphosphate, nicotinamide adenine dinucleotide, DNA and RNA,[53] however, some other DNA and RNA binding proteins rely instead on specialised structures for this purpose, rather than electrostatics. Several conformations have been identified in numerous proteins that appear to be specially adapted to recognise DNA,[54] most notably the helix-turn-helix motif that possesses a helix diameter exactly matched to the width of the major groove in B-form DNA.[55]

In the same journal issue as their α-helix paper, Pauling and co-workers also published the first work on pleated β-sheets. These, they observed, were extended networks of hydrogen bonded peptide chains, but this time the hydrogen bonds were intermolecular, rather than intramolecular, causing a sheet-like array of chains running either parallel or anti-parallel to each other.[56] β-Sheets are made up of β-

strands, which are roughly the equivalent of a helix with two residues per turn. The pleated nature of β-sheets is derived from the sp^3 bond angle of the side-chain bearing carbon causing neighbouring carbon atoms to be alternately above and below the plane of the sheet. It is thought that these type of structures contain between 20 – 28 % of all residues contained within globular proteins.[57]

The β-sheets known today have those two classes, parallel and anti-parallel[58] (see *fig 15*). Anti-parallel sheets are thought to be more stable than parallel as parallel are usually found in the hydrophobic core of proteins and are less common.[59] Anti-parallel β-sheets can be made up of as few as two strands and generally each strand contains six amino acid residues with a repeat dipeptide distance of approximately 7 Å. They are slightly less pleated than their parallel counterparts as the chains are marginally more extended. The inter-strand hydrogen bonding pattern found in anti-parallel sheets consists of alternating 10- and 14-membered H-bond rings and the peptide dipoles running in an anti-parallel fashion cancel each other. Parallel sheets contain five or more strands and are slightly more uniform, with a consistent 12-membered H-bond ring structure, 6.5 Å repeat unit and an overall dipole running towards the N-terminus of the chains.[60] Inter-chain distance in β-sheets is approximately 5 Å and H-bonds are 0.1 Å shorter than those of an α-helix. The H-bonds are also more linear by 3 – 4 °; the average N – H – O bond angle is 160 ° in antiparallel sheets, 161 ° in parallel sheets and 157 ° in an α-helix.[61] β-sheets have a lower potential energy than α-helices, however, α-helices can be considered to have an entropic advantage due to the intramolecular nature of the H-bonds.[62]

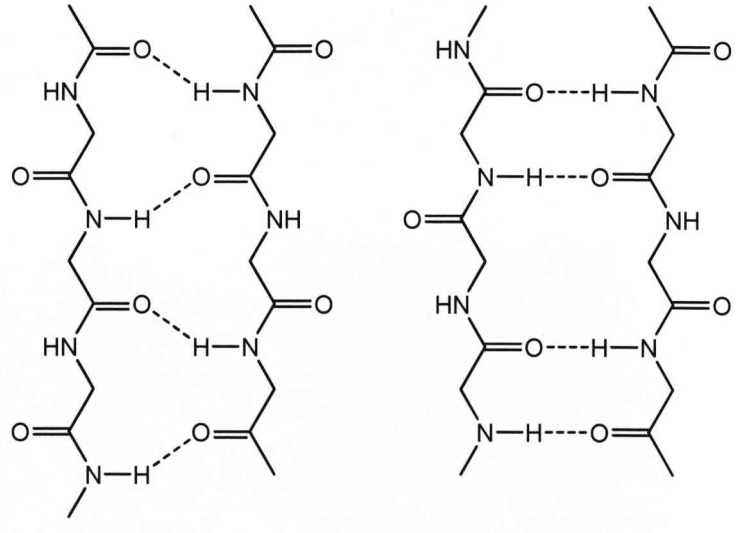


Figure 15: *Parallel (left) and anti-parallel (right) β-sheet hydrogen bonding.*

Alongside their structural function, one of the most important properties of β-sheets is their amphiphilic nature. Side chains are alternately arranged on opposite faces of a β-sheet due to the *trans* nature of the peptide bond. Thus, sheets containing alternating hydrophobic and hydrophilic residues present these side chains on opposite sides of the sheet, giving faces with either hydrophobic or hydrophilic preferences (see *fig. 16*).[63]

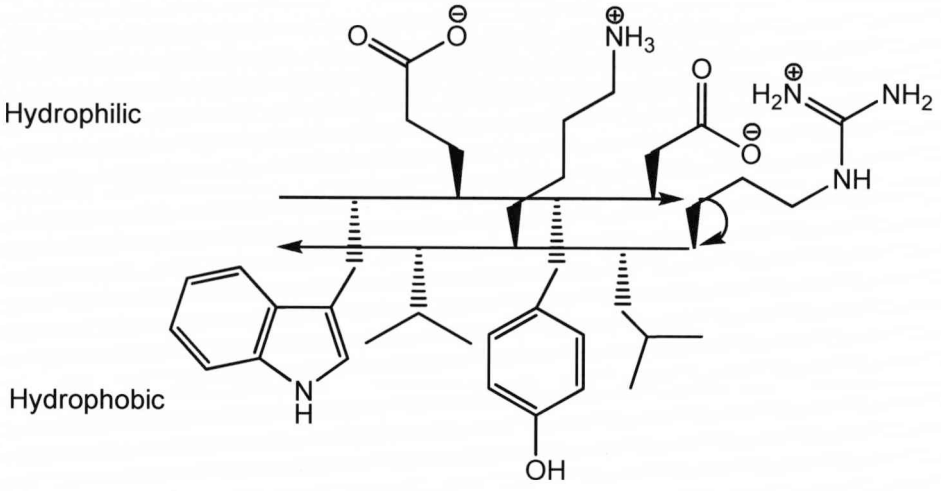


Figure 16: *Representation of an amphiphilic β-sheet at physiological pH.*

The result is that β-sheets are often found at the surfaces of proteins to act as an interface between aqueous and lipid-rich environments; a notable example of which is porins.[64] Porins are also known as "beta barrel" proteins as they are constructed from a large β-sheet that forms a cylindrical tube somewhat reminiscent of a barrel. These structures are often found as membrane proteins that can control the flow of small molecules across a cell, chloroplast or mitochondrial membrane.[65]

With natural systems providing such a richness of intricate and functional secondary scaffolds it is not surprising that much effort has been invested into attempting to mimic the properties of both peptides and oligonucleotides. An enormous array of peptide and nucleotide mimetics can be found in the literature and those that are found to adopt secondary structure have been termed "foldamers".[66,67]

1.5 Secondary Structure in β-Peptides

β-Peptides are amongst the most intensively studied of the foldamer family because of their close relation to natural α-peptides. β-Peptides are sequences of β-amino acids; these have an additional carbon atom between the amine and carboxyl

groups compared to their α-counterparts (see *fig. 17*). The only common, naturally occurring β-amino acid is β-alanine which is found in mammals as a degradation product of uracil[68] and also as a component of the dipeptide carnosine.[69]

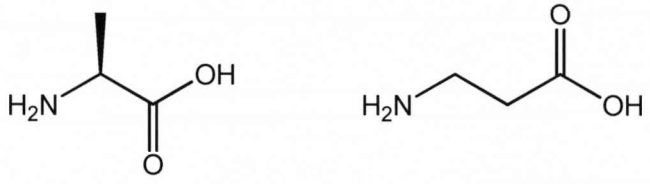


Figure 17: *α-Alanine (left) and β-alanine (right).*

The most attractive aspect of β-peptides, when considering possible therapeutic applications of peptidomimetics, is stability towards cellular enzymes, in particular, peptidases.[70] Most peptidases only recognise α-peptides and thus β-peptides have good resistance both *in vitro* and *in vivo*.[71] It has recently been shown that substituting even a single α-amino acid with a β-residue in a model peptide antigen confers this type of stability without the loss of activity.[72]

Secondary structure was first identified in β-peptides as early as 1965 by Kovacs and co-workers who suggested poly-β_3-*L*-aspartic acid adopted a 14-helix.[73] The term 14-helix refers to the number of atoms in the hydrogen-bonded ring of a helix (see *fig. 18*). Some doubt over the interpretation of the experimental data led to confusion over the exact conformation and this also affected the findings of Yuki and co-workers, who eventually assigned the structure of nylon-3 derivative poly-β-(α-isobutyl-*L*-aspartate) as a β-sheet like formation instead of a helix[74] although their original theory of a helix was later shown to be true.[75] The strong preference of β-peptides for formation of helical structures was demonstrated by examination of nearest neighbour interactions in derivatives of β-alanine and γ-butyric acid by FT-IR. One of the reasons for folding occurring in α-peptides is the predominance of hydrogen bonding between donors and acceptors that are not adjacent, over interactions of residues with their nearest neighbours.[76]

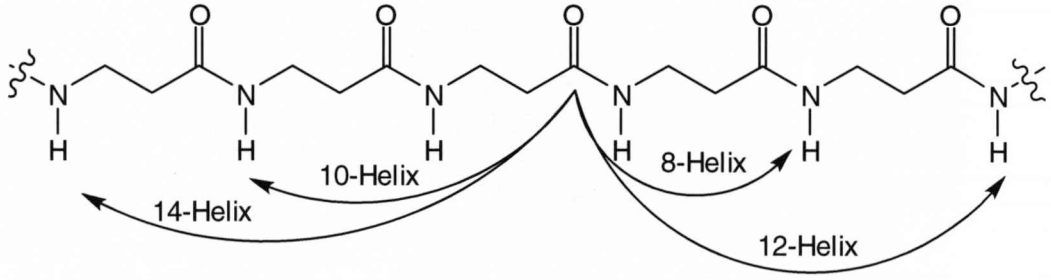


Figure 18: *Hydrogen bonding patterns in β-peptides.*

Gellman asserted that the lack of nearest neighbour connections between sequential H-bond donors and acceptors in β-alanine derivatives indicated that long-range conformational order would be preferred over local, short range hydrogen bonding as H-bonds are stronger when the N-H-O bond angle is closer to 180 °.[77] In essence, this meant β-peptides should favour the formation of hydrogen-bonded rings containing 10 or more atoms over those with just 6 or 8 atoms (see *fig. 19*). γ-Butyric acid derivatives, on the other hand, did show evidence of hydrogen bonding between nearest neighbour donors and acceptors and, therefore, these interactions were thought likely to predominate over any form of higher order folding.

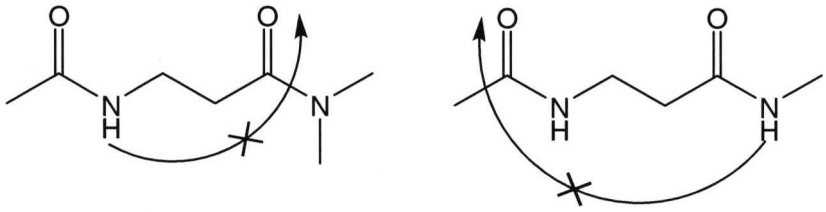


Figure 19: *Disfavoured 6 and 8 atom nearest neighbour H-bonds.*

Pleated sheets, turns, helices and other structural motifs have all since been observed in β-peptides.[78] The most significant difference between α- and β-derived oligomers is the variety of different substructures assumed by β-peptides within these classes and their relative stabilities compared to the corresponding α-analogues. Taking the example of helices, the α-helix is the dominant form of helix in α-peptides; however, depending on their constitution, β-peptides have been observed to form helices based on hydrogen-bonded rings containing between 8 to 14 atoms[66] and remarkably these helices have been shown to be significantly more stable than their α-analogues.[79,80] A possible contributing factor to this increase in stability is a hydrophobic effect, where organic side chains of the amino acids shield the hydrogen-bonded core from solvent interactions thus preventing the break-up of H-bonds by polar solvents.[80] To some extent this effect should also be present in α-peptides[81] and this alone, therefore, is not adequate to explain the increased stability.

As a result of the extra methylene group there are two possible substitution patterns of β-amino acids; β^2 and β^3. β^2 refers to those amino acids with the side chain on the carbon adjacent to the carbonyl whereas β^3 have the side chain attached to the same carbon as the amine (see *fig 20*).

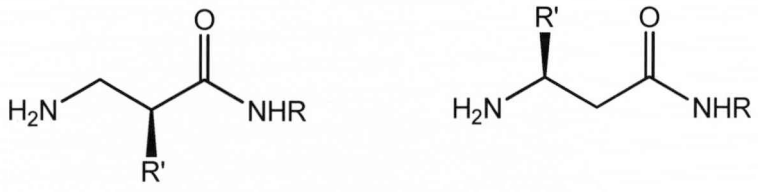


Figure 20: *β^2- (left) and β^3- (right) substituted β-amino acids.*

Folded β-peptide conformations require a *gauche* conformation at the C2 – C3 bond around torsion angle (θ) (see *fig. 21*). A significant contribution to helix stability is made by the preference of β-amino acids for such conformations, thought to be derived from an intra-residue electrostatic interaction between partial charges of the amide nitrogen and carbonyl.[82] The result is a favourable steric arrangement in which the R' group is oriented away from the extended chain.

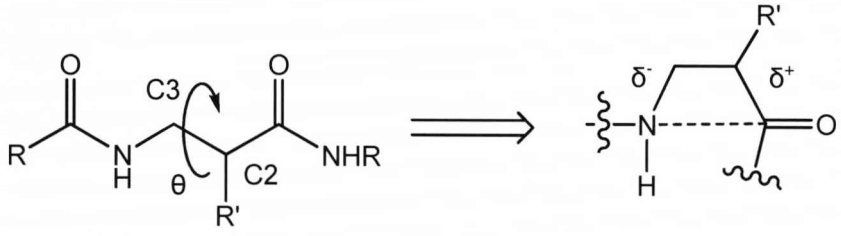


Figure 21: *Intra-residue electrostatic interaction in β-peptides.*

The more highly substituted the amino acid, the larger its preference for adopting the *gauche* conformation, except in the case of *syn*-β^2-β^3-disubstituted residues which favour a *trans* conformation.[78]

β-Peptides containing proteinogenic side chains such as lysine, valine, leucine and alanine show preferences for the type of secondary structure adopted by the oligomer. This is defined by the backbone substitution pattern and stereochemical composition.[83] A peptide consisting of six or more homochiral β^2 or β^3 residues will adopt a 14-helical conformation, one containing (*S*)-β^2/β^3 or geminally disubstituted $\beta^{2,2}$ amino acids will prefer 10-membered hydrogen-bonded rings and sections with (*R*)-β^2 and (*S*)-β^3 units form pleated sheets.[84]

At approximately the same time, both Gellman and Seebach produced very different short oligomeric β-peptides which displayed remarkably similar helical conformations. Seebach and co-workers produced acyclic hexapeptide (H-β-HVal-β-HAla-β-HLeu)$_2$-OH and tripeptide Boc-β-HVal-β-HAla-β-HLeu-OMe from the parent α-amino acids by an elegant Arndt-Eistert/fragment coupling approach.[85] The

tripeptide was shown by X-ray crystallography to adopt a β-sheet like conformation in the solid state; the hexapeptide, on the other hand, showed markedly different behaviour. The circular dichroism spectrum of the hexapeptide showed similarities with α-peptides that were known to have β-sheet conformations. However, the fact that this spectrum was independent of concentration or solvent suggested an intra rather than intermolecular hydrogen bonding network. COSY and nOe NMR experiments finally revealed the hexapeptide to adopt a 14-helical arrangement of residues in d_5-pyridine.

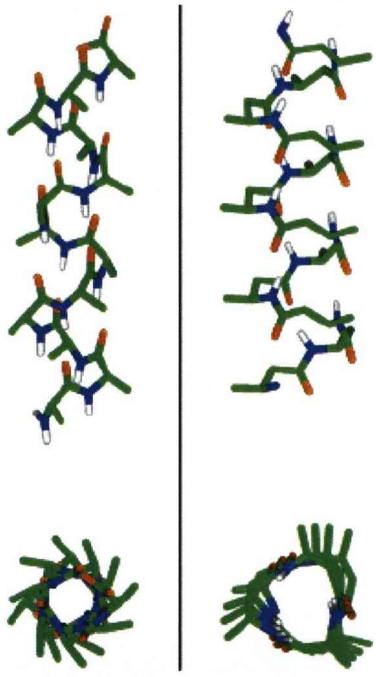


Figure 22: *Comparison of an α-peptide α-helix (left) and a β-peptide 14-helix (right) (taken from reference 79).*

This was considered a somewhat surprising result given that the accepted minimum number of α-residues required to form a secondary structure is approximately 15.[86] It was also thought that the extra flexibility imparted by the additional methylene group in a β-amino acid would mean that even more than 15 β-residues would be required to achieve the same effect; a theory supported by the finding that poly-β-*L*-alanine was known to be unordered in solution.[87] In contrast to the α-helix, the 14-helix hydrogen bonding pattern is n, n + 2 with the dipole running in the opposite direction towards the *C*-terminus. As a result there are 3 residues per helix turn[83] and, unlike the α-helix where the amino acids side chains are approximately evenly distributed around the helical core, the side chains of every third residue in a 14-helix reside almost on top of each other (see *fig. 22*).

Gellman *et al.* also shared the view that longer oligomers would be required in order to generate folding tendencies in the more flexible β-peptides. This is reflected in their use of carbocycles in an attempt to constrain the number of degrees of freedom and induce structure in shorter chains. By modelling helices based on β-alanine, Gellman and co-workers suggested that incorporating the α- and β- carbons of β-alanine into a cyclohexane ring would produce a potentially stable helix by restricting the C2 – C3 bond to a *gauche* configuration. This hypothesis proved to be correct when, upon oligomerisation of *trans*-2-aminocyclohexanecarboxylic acid (*trans*-ACHC, see *fig. 23*), a stable 14-helix was observed in methanol solution by both NMR and isotope exchange experiments, in peptides as short as 6 residues.[88]

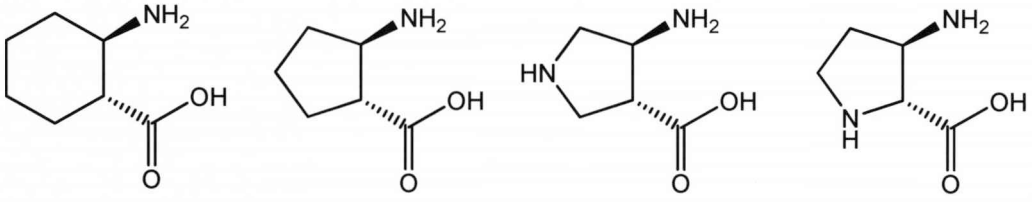


Figure 23: *trans-(ACHC) (far left), trans-(ACPC) (centre left), APC (centre right) and AP (far right) β-amino acids.*

By reducing the size of the carbocycle from cyclohexane to cyclopentane a similar propensity for helix formation was found in *trans*-2-aminocyclopentanecarboxylic acid (*trans*-ACPC, see *fig. 23*). In this instance a 12-helix with n, n + 3 hydrogen bonding, 2.5 residues per turn and the same polarity as an α-helix was observed in equally short oligomers in d_5-pyridine solution[89] and this illustrated the extraordinary degree to which β-peptide secondary structure could be controlled by minor changes to the backbone. ACPC oligomers also highlighted a major drawback suffered by many β-peptide compounds; poor aqueous solubility.[90] In an attempt to circumvent the problems posed by insufficient solubility, inclusion of β-amino acids such as β-lysine and β-glutamate suggested that, unlike their α-counterparts, both oligomers containing cyclic and acyclic residues could retain helical conformations in the presence of water.[90,91] Another method for increasing the aqueous solubility was to substitute the carbocycle of *trans*-(ACPC) with a heterocyclic pyrrolidine analogue. Both *trans*-aminopyrrolidine-4-carboxylic acid (APC) and (2*R*,3*R*)-aminoproline (AP) (see *fig.23*) have been incorporated into β-peptides and confer water solubility whilst retaining the propensity for helix formation.[92]

Since the initial identification of helical tendencies in β-peptides, countless examples of different scaffolds consisting of a wide variety of β-amino acids have been produced. Extending the cyclic theme of Gellman, both Fleet *et al.* and Seebach and co-workers observed helical conformations in compounds containing 4- and 3-membered rings respectively (see *fig. 24*). Oligomers of *cis*-oxetane derived amino acids (see *fig. 24*) displayed a left-handed, 10-helix when analysed by NMR in deuterated chloroform or benzene.[93] 10-membered hydrogen-bonded rings were estimated to be the limit for stable structure formation in β-peptides as nearest neighbour hydrogen bonding interactions were considered unfavourable.[77] This made the emergence of 8-helices an unexpected discovery.

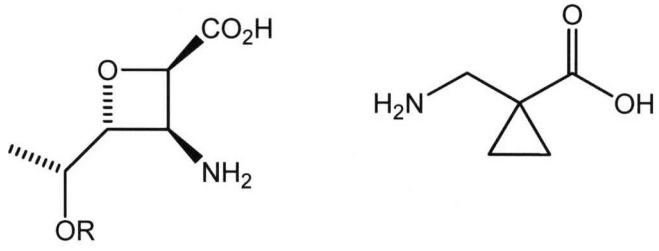


Figure 24: *Oxetane (left) and cyclopropane (right) derived β-amino acids.*

Peptide chains containing a cyclopropane motif (see *fig. 24*) were shown to adopt a "ribbon-like" arrangement of 8-membered hydrogen-bonded rings.[94] Although it was a deviation from the strategy of constraining the C2 – C3 bond to *gauche*, the amino acid being essentially $\beta^{2,2}$ disubstituted, the limitation in this instance was the preference of the C=O – cyclopropane bond for the *s-cis* conformation. It is interesting to note that the peptides adopted a 10-helical shape when the cyclopropane ring was replaced with a cyclohexane. 8-Helices have since been characterised in numerous oligomers, each appearing to demonstrate that the key to smaller hydrogen-bonded ring sizes is a steric or electronic demand favouring nearest neighbour over longer range interactions.[95,96] Homo-oligomers almost always produce regular, easily classified structures with a single identifying repeating hydrogen bonding unit. However, mixed sequence oligomers with unlike residues often produce unexpected and unconventional results.

"Mixed helices" are frequently obtained when amino acids with unlike substitution patterns are included in the same oligomer; Characteristic of this class of compounds is the deviation from regular 14- and 12-helices towards structures with alternating and, in some cases, co-existing hydrogen-bonded rings. The most well known example of this is the 12/10/12 helix[97] and this was discovered during

the study of a series of β-peptides derived from β-alanine. These peptides possessed leucine and valine-like side chains in an alternating β^2 and β^3 substitution pattern. Rather than being intrinsically favoured, molecular mechanics calculations showed formation of this atypical helix resulted from steric interactions of aliphatic side chains de-stabilising conventional helices in favour of the 12/10/12 pattern.[98]

Whilst the mixed 12/10/12-helix was constructed with exclusively β-amino acids, other more recent examples of mixed helices have featured combinations of α/β amino acids as well as a variety of non-peptidic residues.[99] Similarly remarkable structures have been characterised within these peptides. Oligomers made up of *trans*-ACPC and α-amino*iso*butyric acid were shown to adopt either an 11- or mixed 14/15-helix by crystallography and the exact conformation adopted was highly sequence dependent.[99,100] Mixed 11/9-helices were characterised by Srinivasulu and co-workers in peptides consisting of alternating α-*D*-phenylalanine and β-H-valine by NMR. Baldauf *et al* have shown by a series of *ab initio* calculations that the intrinsic flexibility of γ-amino acids may afford easy access to 12/10-, 18/20-, 11/13-, and 20/22-helices.[101]

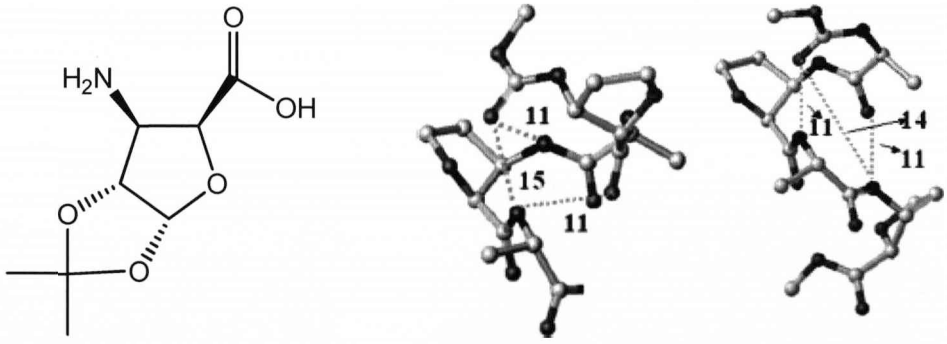
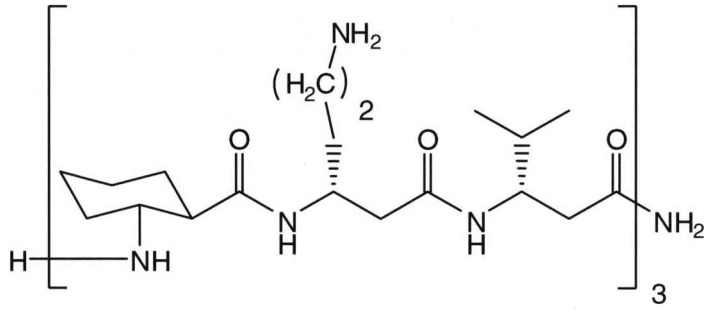


Figure 25: *cis-β-furanoid sugar amino acid (left) and representations of mixed 11-, 14-, and 15-helices (right, taken from reference 103).*

These calculations were borne out when Sharma and co-workers identified the first example of a 12/10-helix without β-amino acids in a peptide containing *L*-alanine residues sequenced with γ-amino acids derived from *D*-mannose.[102] An 11/13-helix was also identified in oligomers of *D*-xylose derived β- and γ-residues and these compounds reflect the increasing use of sugars in peptide chemistry. Possibly the most unusual example of a mixed helix is also found in peptides containing carbohydrate derived amino acid residues. Jagadeesh *et al* produced peptides with alternating *cis*-β-furanoid sugar amino acid and *L*-alanine residues and found coexisting 11-, 14- and 15-helices by NMR (see *fig. 25*).[103]

When considered as a whole, all of these results demonstrate that β-peptides have the potential to be tuned to adopt specific shapes, possibly for specific applications, by altering their constitution. Currently, β-peptides have a wide variety of applications including peptidomimetics[104] and nanomaterials.[105] Significantly, β-peptides, such as those shown in *figure 26*, have also demonstrated anti-microbial activity against both Gram positive and negative organisms.[106]



β^{3}HPhe - β^{2}HVal - β^{3}HPhe - β^{2}HVal - β^{3}HLys - β^{2}HVal - β^{3}HLys - β^{2}HVal - β^{3}HPhe

Figure 26: *β-peptides with anti-microbial activity.*

It is thought that the mode of action involves disruption of membrane viability by binding to the membrane and forming amphiphilic helices which facilitate diffusion across the cell boundary. Activity against *Staphylococcus aureus*, *Enterococcus faecium*, *Escherichia coli* and other bacteria has been observed, importantly at much lower concentrations than is required for haemolysis of mammalian cells.

1.6 Higher Order Structure in β-Peptides

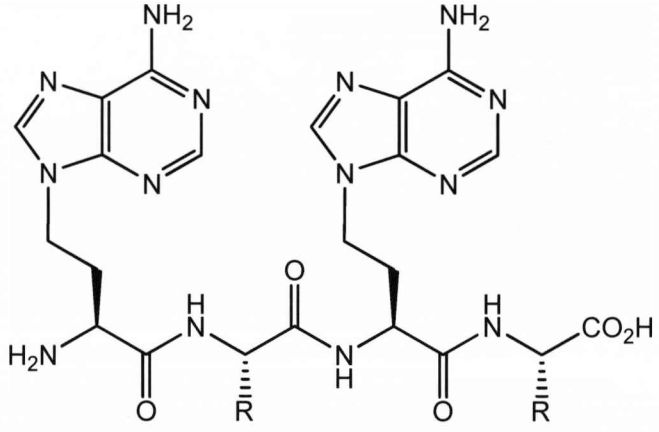
A major goal of the study of synthetic peptides is to enable the construction of artificial proteins. Recently, several groups have taken significant steps towards demonstrating that "β-proteins" are a realistically attainable target and have produced ordered assemblies of β-peptide helices resembling naturally occurring α-helix bundles, as well as semi-synthetic enzymes containing β-peptide components. Arnold and co-workers replaced Asn113 and Pro114 in ribonuclease A with an di-nipecotic acid unit resulting in a gain in conformational stability whilst enzymatic activity was fully retained.[107] Gellman and co-workers have identified helix bundle structures in oligomers containing alternating α- and β-amino acid residues[108] whilst Schepartz *et al.* observed an octameric bundle of parallel 12-residue β-peptide helices containing exclusively β^3-residues. This bundle displayed similar properties to an α-helix bundle, such as a hydrophobic core.[109] This structure, however, was only observed at high concentrations of the β-peptide as the helices were relatively

poor at self-recognition. A more stable analogous structure was produced when the 12-residue peptide was bound together by 4 β-homoalanine residues into a 28-residue anti-parallel helical dimer which then associated into a tetrameric bundle-like arrangement.[110]

These examples clearly demonstrate the potential for mimics of natural biopolymers to possess extremely desirable properties. By further functionalisation and modification of peptide analogues, enormous volumes of novel compounds are now under development and more useful properties and specialised applications are being both sought and discovered.

1.7 β-Peptides with Nucleobase Modifications

Another method of introducing ordered structure to peptide analogues is to utilise the predictable associative properties of nucleobases in order to bind structural units together. Peptides modified to contain nucleobases as side chains are the oldest class of peptide/oligonucleotide hybrid compounds; the synthesis of nucleobase derived alanine dates back to 1969.[111] Although much of the work regarding peptide backbones with nucleobase moieties is now focussed on the oligonucleotide binding peptide nucleic acid (PNA, see *section 1.8*), there are several examples of true polypeptides functionalised with nucleobases.



R = H, CH_2OH, $CH(OH)CH_3$ or $CH_2C_6H_5OH$

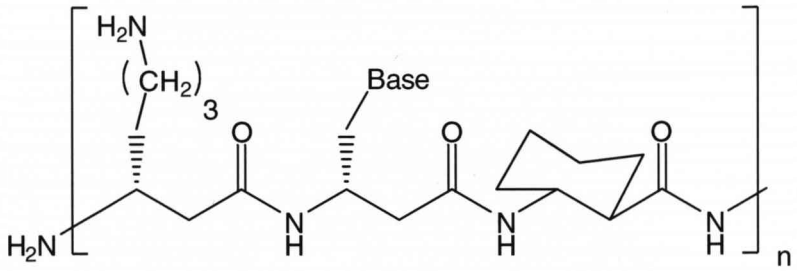
Figure 27: *Triplex forming tetrapeptide.*

Many of these oligomers have proved unsuccessful at binding to complementary oligonucleotides, but not surprisingly the most successful oligonucleotide mimics are those which closely match the physical dimensions of an oligonucleotide. Yamazaki and co-workers reported triplex formation by a series of

tetrapeptides consisting of glycine, serine, threonine and tyrosine residues alternately sequenced with an *L*-α-amino acid unit functionalised with adenine (see *fig. 27*). The spacing between the pair of heterocyclic bases was approximately equal to that found in oligonucleotides and, even though this was a particularly short peptide, the binding affinity for both poly d(T) and poly d(U) sequences was significantly higher than that of a natural d(A) dinucleotide.[112] This was thought to be due to lack of electrostatic repulsion between the neutral polypeptide and anionic oligonucleotide.

As it became more apparent that conformational restriction of the backbone was an important factor in improving the stability of secondary structure in β-peptides, interest in α-peptide hybrids naturally shifted to more rigid amino acids. The 5-membered ring motif, which had so successfully provided a platform for the formation of secondary structure in β-peptides, was introduced in the form of proline[113] and oligomers containing hybrid proline/nucleobase monomers were eventually to evolve into pyrrolidine-type PNA (see *section 1.8*). However, β-peptides containing nucleobases which were not designed to recognise oligonucleotides are increasingly being used as tools for assembly of complex structures.

The predictability of both association of nucleobases and the helicity of β-peptides is a powerful combination for the *de novo* design of molecular scaffolds. 14-helices in β^3-peptides are arranged with 3 residues per helix turn and Diederichsen *et al.* demonstrated that placing a nucleobase side chain in an n, n + 3 pattern resulted in one side of the helix being exclusively occupied by the bases[114] (see *fig. 28*). The strands were water solubilised by the inclusion of lysine residues. Antiparallel duplex formation via Watson-Crick base pairing was then observed for strands containing 4 homologous bases between A-T, T-T, C-C^+, T-C and T-G pairs as determined by UV melting experiments.[115] The lack of specific interactions between homo-G and homo-C strands was a surprising result given that a self-recognising GCGC sequences had C-G pairs with the highest melting temperature (*Tm* = 80 °C) of any of the base pairs tested so far.[116] Two conclusions can, therefore, be drawn from these results. Firstly, nucleobase association is a useful tool in the construction of higher architectures from helical β-peptides and secondly, the molecular design is yet to reach an optimal level as aggregation is still prevalent in some sequences where Watson-Crick pairing is also possible.



Base = A, T, C, G

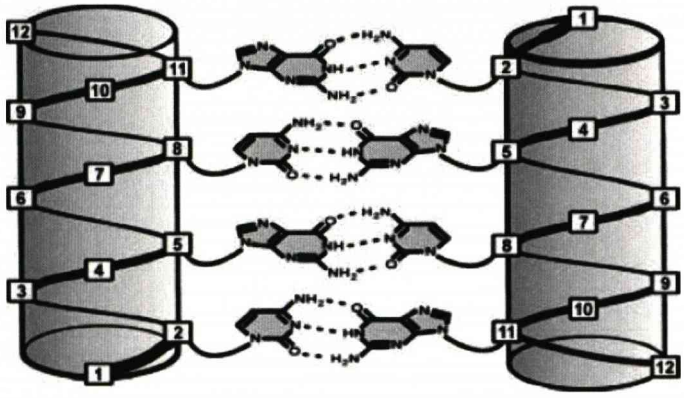


Figure 28: *Nucleobase functionalised β-peptide (top) and a representation of duplex formation (bottom, taken from reference 116).*

1.8 Peptide Nucleic Acids

The study of peptides functionalised with nucleobases has led to the discovery of a remarkable class of oligonucleotide analogues known as peptide nucleic acids (PNA). Peptides bearing a nucleobase side chain can be regarded as the chemical ancestors of PNA; indeed some PNA strands are still produced with β-peptide residues,[117] however there are some key differences between PNA and a "true" peptide. PNA consists of nucleobases bound to a neutral pseudo-peptide backbone via an acyl linker and was designed as a sequence specific DNA duplex recognition agent by Nielsen and co-workers in 1991.[118] It was hoped that PNA would be able to mimic single strands of DNA that can bind to duplexes via Hoogsteen hydrogen bonding in the major groove to form a triplex. However, PNA was later found to be an extremely effective structural mimetic of DNA and RNA and posses unusual properties of its own.[119] As a consequence of very different chemical and physical properties to natural oligonucleotides PNA has, over the last 15 years, attracted much attention in the area of antisense and antigene therapy.[120] Unlike DNA, PNA is acid stable and will not suffer cleavage of the glycosylic bond

even under the most harshly acidic conditions.[121] PNA is also stable to mildly basic environments but will undergo *N*-acyl transfer reactions in the presence of a strong base[122] (see *figure 29*). PNA can, therefore, be synthesised by both Boc and Fmoc solid-phase methods.

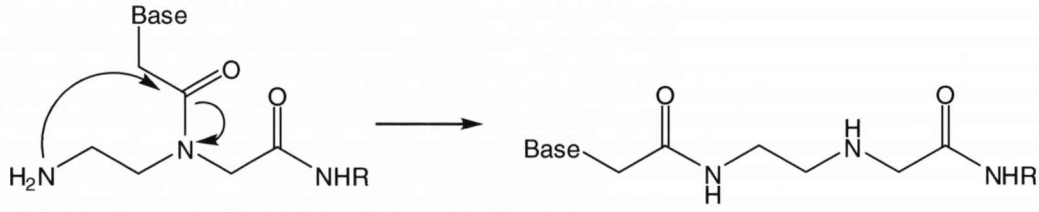


Figure 29: *N-acyl transfer reaction in PNA.*

In terms of biological stability, PNA is attractive as a potential drug candidate because of its resistance to both peptidase and nuclease enzymes under conditions that result in total degradation of unmodified control oligomers. Both homologous and mixed sequence PNAs have been shown to be stable to human and other mammalian peptidases as well as those found in fungi and bacteria such as *E-coli*.[123] Similar sequences also show no decomposition when incubated with S1 nuclease or DNase[124] and for these reasons PNA is considered ideal for *in vivo* therapeutic applications.

One of the problems associated with PNA for such applications is solubility.[125] As the pseudo-peptide backbone bears no charge, unlike an oligonucleotide, and possesses no hydrophilic side chains as natural peptides do, aqueous solubility is often very poor and decreases with increasing chain length. PNA duplexes remain viable in concentrations of up to 70 % of an organic co-solvent such as DMF,[126] illustrating the stark contrast in solubility properties compared to DNA or RNA. A common strategy to improve the solubility, and hence bioavailability, of PNA is incorporation of natural hydrophilic amino acids, especially lysine, as additional side chains or as modifications to the pseudo-peptide backbone.[127] As the primary amine of the lysine side chain has a pKa of 10.5, it is usually protonated in aqueous solution at pH 7 and thus helps to improve the solubility of the otherwise uncharged peptide chain.

The archetypal PNA consists of an *N*-(2-aminoethyl)glycine backbone which is bound to the nucleobase at the amino moiety via an acetyl linker[118] (see *fig. 30*). Monomers containing the four DNA nucleobases were synthesised by varying routes,[128] then the PNA was assembled using standard BOC protocols.[121] This

provided a good basis for duplex recognition as it mirrored the bonding pattern seen in oligonucleotides; 6 bonds between neighbouring nucleotides and 3 bonds separating the backbone and heterocyclic base. This became known as the "6 + 3 rule".

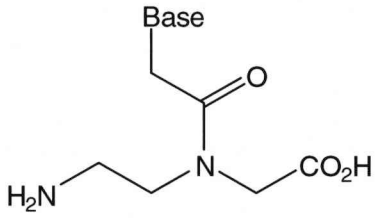


Figure 30: *N-(2-aminoethyl)glycine monomer unit of PNA.*

PNA has been shown to anneal with single stranded DNA and RNA to form helical structures which are similar to unmodified duplexes (see *fig. 31*), self recognise to form PNA/PNA duplexes and will also bind duplex DNA as it was originally designed to do. The complexes formed as a result of the latter two interactions have some unusual and distinctive characteristics. PNA/DNA duplexes are the least stable of the hybrid duplexes formed by PNA followed by PNA/RNA and PNA/PNA duplexes are the most stable.[129] The strands are arranged anti-parallel; that is the amino terminus of the PNA strand is opposite the 3'- end of the oligonucleotide. In contrast to natural oligonucleotides, parallel arrangements of PNA/DNA and PNA/RNA duplexes have also been observed but these are less stable than the anti-parallel arrangements[130] with very much slower kinetics of formation.[131]

The DNA strand of a PNA/DNA duplex resembles the conformation found in B-form DNA and the sugar pucker of the deoxyribose ring is retained as C2'-*endo* but the helix also contains elements of the RNA like A-form.[132] The helix is right-handed with a slightly larger pitch, 13 base pairs per turn compared to 10 in B-form DNA, and the major and minor grooves are wider and narrower than B-form DNA respectively. The amides of the PNA backbone are all in a *trans* conformation and are mostly orientated towards the solvent. The amides of the nucleobase linkers are aligned towards the carboxy terminus of the pseudo-peptide but interestingly there is no hydrogen bonding present between any amide and carbonyl groups.[133] PNA/RNA duplexes are also a composite of A-form and B-form helices. In this case the backbone amide bonds switch to *cis* stereochemistry and, in common with PNA/DNA duplexes, the ribose sugar retains the C3'-*endo* pucker seen in normal RNA duplexes. No hydrogen bonding is observed within the PNA strand. A

significant contribution to the stability of annealed PNA hybrid duplex structures is made by the neutral nature of the peptide based backbone. In DNA and RNA duplexes there is a repulsive electrostatic effect between the two anionic strands; however, the lack of charge on the PNA oligomer eliminates this, thereby helping to decrease the overall energy of the helix.[134]

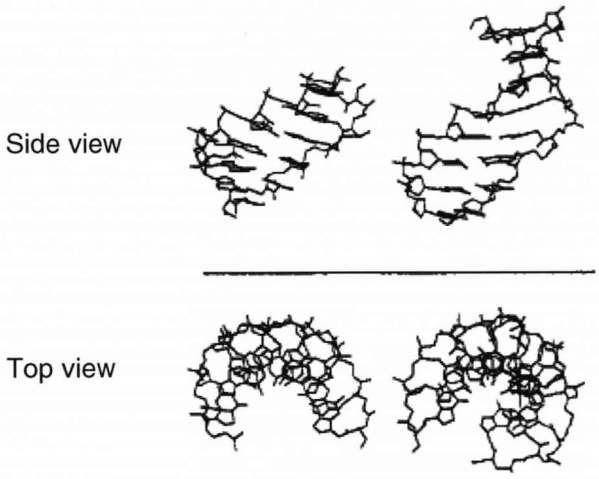


Figure 31: *Representations of a PNA/RNA duplex (left) and a PNA/DNA duplex (right) (taken from reference 132).*

The PNA/PNA duplex forms a unique helix which is known as "P-form"[135] (see *fig. 32*). Like the analogous duplexes seen with DNA and RNA, the PNA duplex shares characteristics with both the A- and B-form helix and also possesses some individual features.

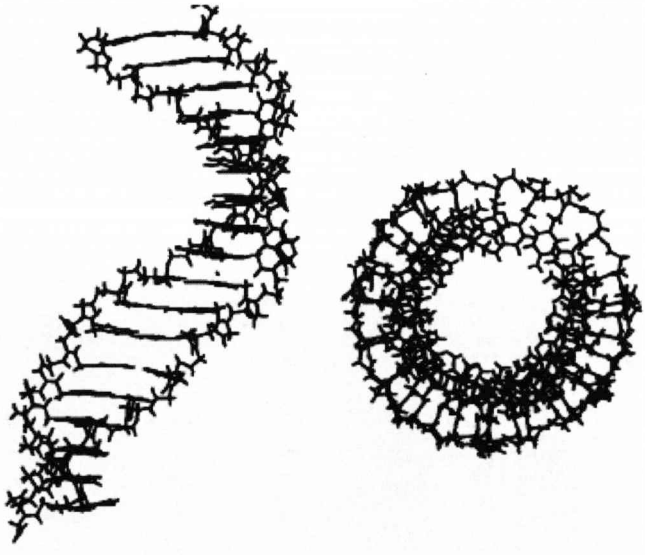


Figure 32: *Side and top views of a P-form duplex (taken from reference 135).*

As is evident from the deviation of PNA/DNA/RNA hybrid duplexes from standard models of A-form and B-form helices, ethylaminoglycine based PNA's are not ideal DNA or RNA binders. As a consequence of this there are now a multitude of modified PNA's made from a wide variety of monomers (see *fig. 33*). By violating the 6 + 3 rule, for example by extending the linker to the heterocyclic base or extending the backbone chain[136], many of these PNA's possess inferior *Tm's* to the original ethylaminoglycine PNA by as much as 22 °C.[137] However, there are a few notable examples which both adhere to and break this rule and show similar if not improved hybridisation stability.

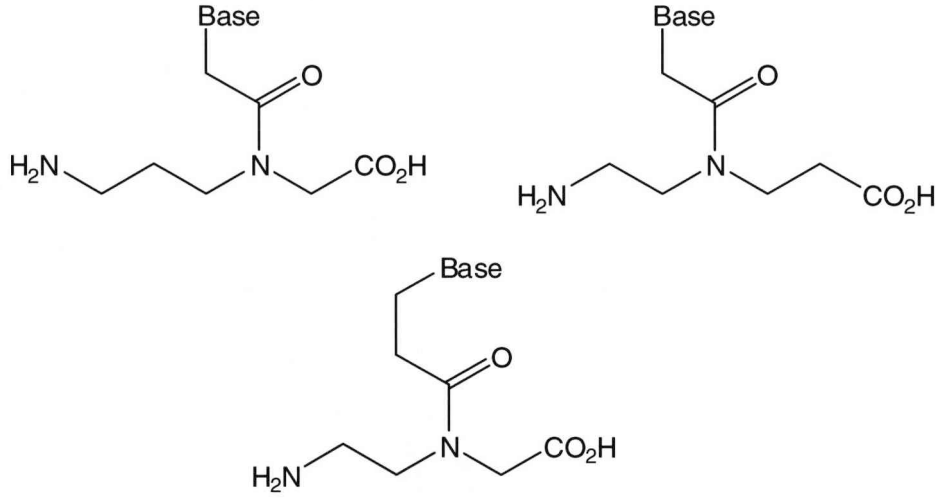


Figure 33: *Modified PNA monomers.*

In energetic terms, helix formation is a large gain in enthalpy and a large drop in entropy. The loss of entropy is due to the formation of order within the extended structure. Conformational constraint of the ethylene linker to a gauche configuration around the C-C bond,[132] much like that seen in β-peptides (see *section 1.5*), ensures that this loss of entropy is not large enough to be prohibitive for helix formation. This led to the hypothesis that further limitation of the flexibility of the pseudo-peptide would further increase the already impressive stability of PNA/DNA/RNA hybrid duplexes. In a similar manner to modifications made to β-peptides, some of the most successful of these alterations to the backbone have come with the introduction of cyclic motifs. Pyrrolidine,[138] aminoproline,[139] cyclohexane[140] and cyclopentane[141] sub-units have all been incorporated into PNA backbones (see *fig. 34*) with varying degrees of success. The pyrrolidine and cyclohexane motifs also demonstrate that stereochemistry has an influence on the binding properties of the oligomer.

Cyclohexyl PNA binds to both DNA and RNA with only a minor loss of duplex stability; however, this is only holds true for the (*S,S*)-isomer. The (*S,S*)-configuration was shown by molecular modelling studies to constrain the backbone to a conformation similar to ethylaminoglycine PNA whereas the (*R,R*)-isomer significantly perturbed this structure and resulted in loss of stability.[140] Pyrrolidine derived PNA, which clearly violates the 6 + 3 rule, showed distinct preferences for binding DNA over RNA, anti-parallel over parallel binding modes and, unusually, hybridisation between self-complimentary strands was not apparent.[138] This indicates that the introduction of more complexity into PNA oligomers imparts certain binding selectivity rules which are highly dependent on factors such as local stereochemistry.

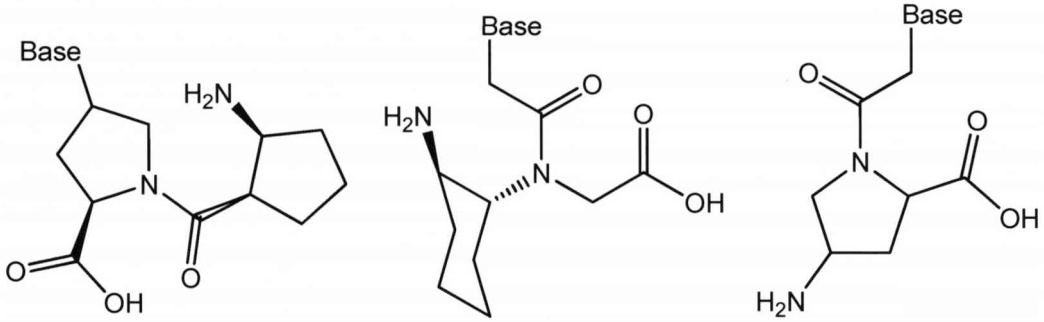


Figure 34: *Cyclic PNA monomers.*

The effect of stereochemistry was more vividly demonstrated by a simple chiral modification to *N*-(2-aminoethyl)glycine PNA's. Introduction of an *L*-serine type side chain at the γ-carbon (see *fig. 35*), whilst also functioning as a solubility enhancer, brought about helical structure, reminiscent of foldamers seen in 14-helical β-peptides, in single stranded oligomers which retained strong binding affinity for both DNA and RNA.[142] This result demonstrates that it should, in theory, be possible to design PNA possessing the conformational constraint thought necessary to afford PNA/DNA/RNA hybrid duplexes with near-natural helicity and hence, produce PNA oligomers with tuneable binding properties.

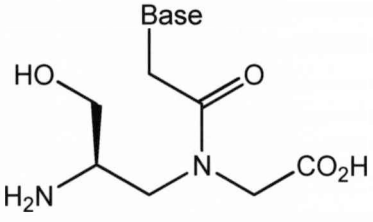


Figure 35: *γ-modified PNA monomer.*

Possibly the most interesting property of PNA is its binding modes during triplex formation. The type of triplex formed by PNA and DNA is highly sequence dependent. For example, PNA oligomers with high proportion of pyrimidines will bind to DNA to afford a $(PNA)_2$/DNA triplex, whereas C-rich PNA hybridised with GC rich DNA prefer $(DNA)_2$/PNA triplexes. The $(PNA)_2$/DNA triplex is extraordinarily stable with a *Tm* for a 10-mer homothymine sequence binding to $d(A\text{-}T)_{10}$ of 73 °C, compared to 23 °C for an unmodified T_{10}.[118] $(PNA)_2$/RNA triplexes are also known to have equally high stabilities and this leads to another remarkable phenomenon; duplex invasion.[143] Under certain circumstances, when targeted to a duplex with an A-T rich region, the tendency to form the $(PNA)_2$/DNA triplex is so strong that the DNA duplex will be invaded, displacing one strand out and forming a triplex inside the host duplex (see *fig. 36*). This is a highly unusual binding mode and is thought to be stabilised in part by hydrogen bonds between the pseudo-peptide amides and the phosphate bridge of the DNA strand.[144]

It is also interesting to note that higher order structures such as G-quadruplexes and the *i*-motif can also be formed by PNA. Both hybrid PNA/DNA/RNA G-quadruplexes[145] and *i*-motifs[146] have been observed and analogous structures can also be formed by PNA alone[147], emphasising the fact that PNA is an extraordinarily good oligonucleotide mimic.

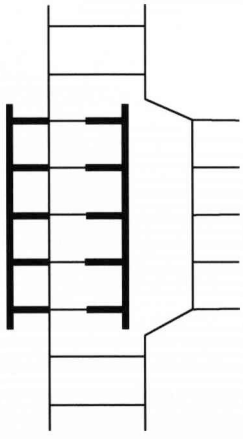


Figure 36: *Schematic of duplex invasion by $(PNA)_2$/DNA triplex, (adapted from reference 132).*

PNA can be thought of as a hybrid between peptides and oligonucleotides in which nucleobases are introduced to a pseudo-peptide backbone. Whilst PNA is structurally similar to a standard peptide it is logically desirable to attempt to incorporate modifications into natural oligonucleotide backbones. One approach is

to replace the phosphate bridge in a standard oligonucleotide with a peptide bond generating a class of oligonucleotide/peptide hybrids which are more structurally related to oligonucleotides than peptides.

1.9 Internucleoside Amide Linkages

As with most analogues of nucleic acids or peptides, there are a multitude of structurally related oligomers bearing an internucleoside amide linkage. The five easily accessible amide isosteres of the natural 3'-5'-phosphodiester unit have been produced (see *fig. 37*) and show interesting properties when incorporated into oligonucleotides and hybridised with complementary DNA or RNA strands. Both isomers of 3'-amides have a destabilising effect on DNA/RNA duplexes in the region of -2 – -4 °C on average per residue where the DNA strand carries the amide modification.[148] This effect is more pronounced when the R group is sterically larger[149] and is also observed to a lesser effect in 5'- analogues.[150]

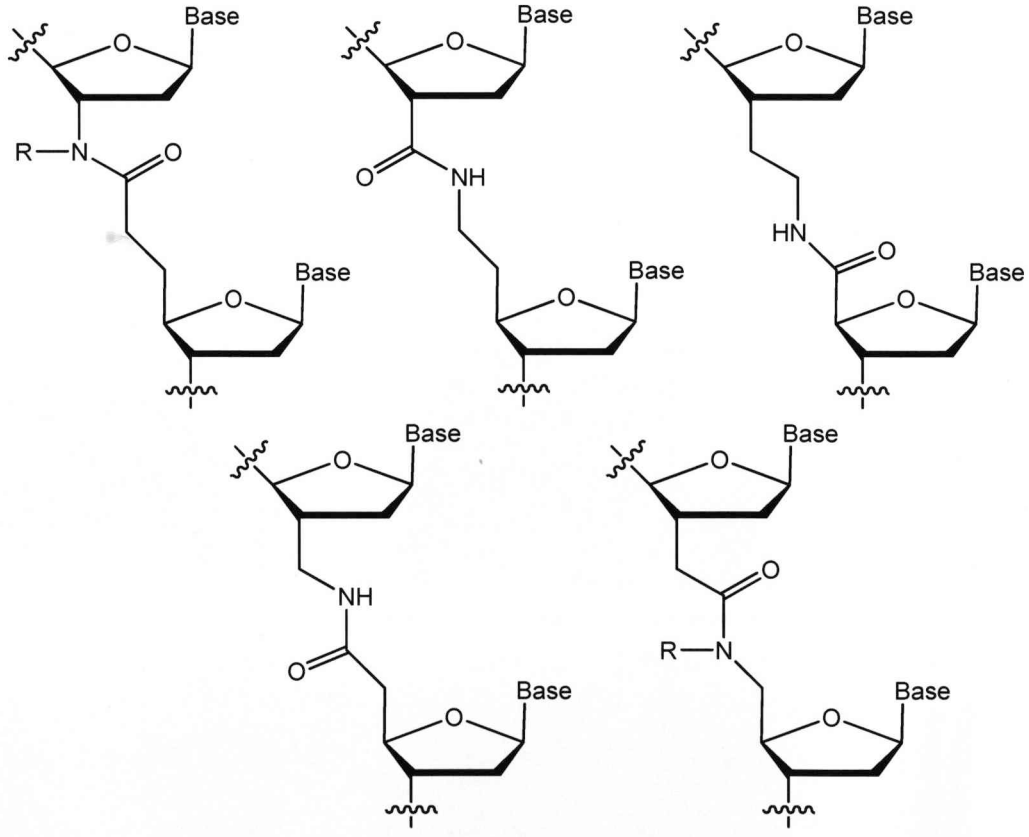


Figure 37: *Five isosteric internucleoside amide linkages.*

It was suggested by DeMesmaeker and co-workers that the destabilisation was brought about by a shift in the preference of the deoxyribose sugar ring to a

rather unusual *C*4'-*exo* or *O*4'-*endo* conformation induced by the structural rigidity of the amide bond. This hypothesis was supported by the observation that amide linkages not directly connected to either of the sugars did not significantly alter the backbone torsion angles or sugar pucker from that of the "ideal" A-form helix. This resulted in either a very minor destabilisation of the duplex or, in some cases, a minor stabilising effect of less than 1 °C.[151] These results were highly sequence dependent and, as with the 3'-amides, increasing the R group size resulted in a larger decrease in duplex stability.[152] Much of the work on hybridisation of amide modified oligonucleotides is focussed on hybridisation with RNA for two reasons; firstly, RNA is often the target for therapeutic applications of such modified oligomers and secondly, modified nucleosides tend to prefer RNA-type sugar puckers and thus form more stable duplexes with RNA. The second factor has been exemplified by the destabilisation of a DNA/DNA duplex by amide linkages by approximately -1.5 °C per residue[152] and was also exemplified in 2' – 5' amide linkages where, although a 12-residue 2' – 5'-phosphate DNA/3' – 5'-phosphate RNA helix was significantly destabilised by around -8 °C, the corresponding 2' – 5' amide DNA/3' – 5'-phosphate DNA duplex stability decreased by more than -12 °C.[153]

One unusual property of modified oligonucleotides bearing an internucleoside amide linkage is their ability to stabilise hybrid DNA duplexes containing a single base pair mismatch. This improved stability is achieved when the amide group is not directly linked to either of the nucleoside sugar rings and confers almost a 3 °C enhancement in duplex stability compared to an unmodified DNA strand.[152] Contrary to previous results, increasing the steric bulk of the R group from H to methyl further stabilised the mismatched strands. This may be an indication that the rigidity of the amide linkage is a stabilising factor by more effectively mimicking an A-form helix, whilst a mismatched DNA nucleotide may prefer to revert to a B-form like conformation, reducing the overall stability of the duplex. Recently, amide modifications have been inserted into RNA strands resulting in the identification of the 3'-$CH_2NHCOCH_2$-5' linkage as a superior stabiliser of RNA/RNA duplexes compared to the isomeric 3'-$CH_2CONHCH_2$-5' by between 1 and 2.4 °C per modification.[154] The reason for this behaviour remains uncertain, but may be due to differential interactions within the hydrogen bonded solvent network.

Amide-type linkages which are not isosteric with a phosphodiester have also been examined as oligonucleotide analogues. An extra atom was introduced in an

attempt to offset the shorter length of the amide bonds when compared to those of the phosphodiester and thereby improve duplex stability.[155] Gait and co-workers showed in early work with acetamidate linkages (see *fig. 38*) that oligomers exclusively joined with this group did not possess the ability to recognise complementary polyribonucleotides[156] probably due to an unfavourable conformation of the backbone. In a similar study, Pallan *et al.* showed that the introduction of a methyl substituent to the acetamidate group facilitated RNA recognition[157]; thus, changing the substituent could assert at least some control over backbone conformation. These duplexes, however, remained considerably less stable than an unmodified control.

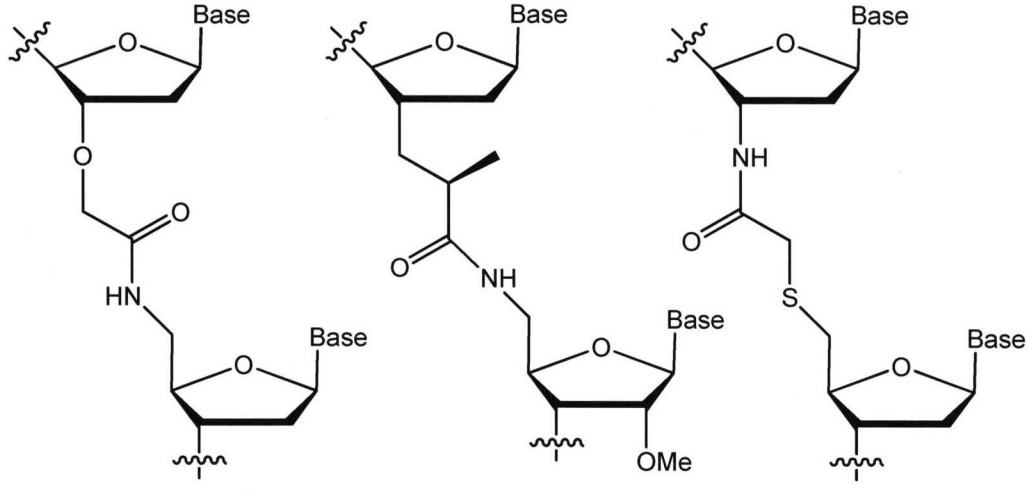


Figure 38: *Non-isosteric amide-type internucleoside linkages.*

The inclusion of a 2'-methoxy group at the 5'-nucleoside of the methyl modified acetamidate brought about stabilisation of the duplex compared to the control. 2'-OMe groups are well known to induce an RNA-like sugar pucker in deoxynucleotides[158] and this indicated that the bulkier methyl bearing linker may tend to favour an RNA-like *C*3'-*endo* conformation in the sugar. This suggests that the most important factor for RNA recognition and duplex stability with amide modified DNA strands is the ability of the linkage to constrain the sugar to a *C*3'-*endo* conformation. The same study demonstrated that switching from an acetamidate to a 5-atom amide linkage by changing 3'-*O* to 3'-*C* whilst retaining the 2'-OMe moiety also produced a duplex stabilising effect; again highlighting that the critical nature of the sugar pucker. It was also shown that imparting a preference on sugar pucker through stereochemical means may be possible, given that the *R* stereoisomer of a methyl substituted amide linkage (see *fig. 38*) was more stabilising than the *S* isomer.[157] Since these findings, various 5-atom amide linkages

have been observed to bind to RNA in a similar manner to that of isosteric modifications. Some were found to improve the stability of duplexes whilst increasing bioavailablity; for example those bearing a thioether modification (see *fig. 38*).[159]

1.10 Project Aims

The primary goal of this project was to produce β-peptides derived from nucleoside β-amino acids and examine the structure of these highly functionalised oligomers (see *fig. 39*). Encouraged by the striking similarity of the proposed oligomers with foldamers produced by Gellman and co-workers from *trans*-ACPC, the objectives set were to synthesise a nucleoside β-amino acid derived from deoxynucleoside thymidine which was compatible with peptide synthesis by standard solid-phase peptide synthesis methodologies. This monomer would then be utilised to synthesise nucleoside β-peptide homo-oligomers of varying length and these, in turn, would be examined for the presence of ordered secondary structure by NMR and, if applicable, other spectroscopic techniques.

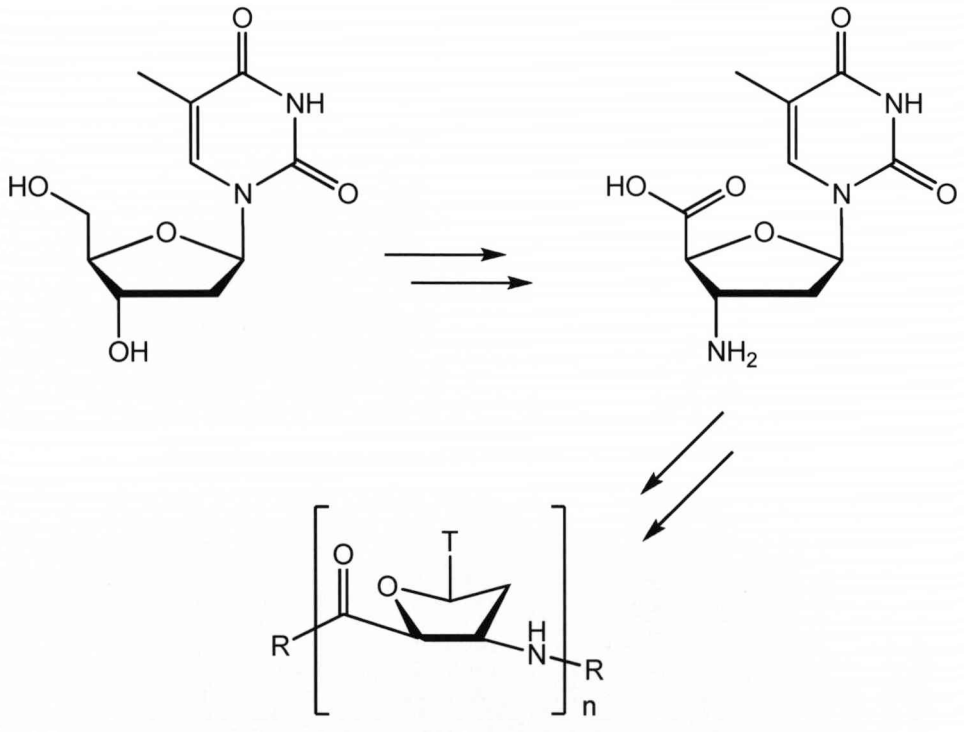


Figure 39

A logical extension of this work was to apply these methods to other deoxynucleosides, particularly 2'-deoxyadenosine, and investigate hybridisation of the thymidine derived β-peptides with complementary β-peptide and natural DNA sequences. It was also desirable to use these unique amino acids to explore the possibility of peptide synthesis by means other than solid-phase chemistry, especially more non-traditional based approaches.

It should be noted that after the commencement and publication of the findings of this work regarding β-peptides derived from thymidine, independent studies on oligomers of the same constitution were published by Chandrasekhar *et al.*[160] The implications of this will be discussed in the results and discussion section (see *section 2.7*)

Chapter 2

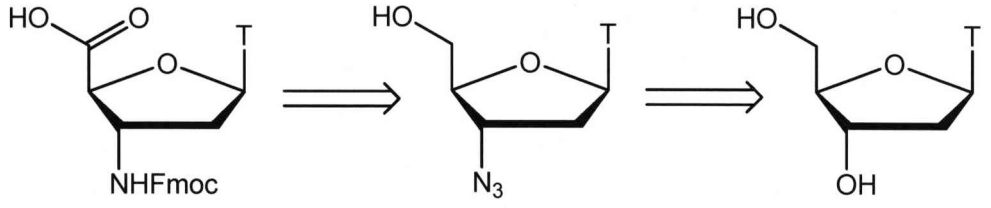
Results and Discussion 1:

Nucleoside β-Amino Acid and β-Peptide Synthesis

2 Results and Discussion

2.1 Thymidine-derived β-Amino Acid Synthesis

Many previous synthetic studies of 3'-amino-3'-deoxythymidine have the corresponding azide as a common intermediate. It was clear that an efficient synthesis of 3'-azido-3'-deoxythymidine (AZT) would be integral to the overall route design (see *fig. 40*). The proposed route to our target protected amino acid (**1**) from AZT (**2**) involved reduction of the 3'-azide group to afford a 3'-amine. This amine could then be protected appropriately in preparation for solid-phase peptide synthesis leaving the last step as oxidation of the 5'-*C* to yield the desired amino acid.[161] The choice of Fmoc as the amine protecting group for SPPS was driven by the sensitivity of the glycosylic bond in 2'-deoxynucleosides to acid catalysed hydrolysis and the desire to avoid such a side reaction occurring during SPPS. Base-labile Fmoc protection would avoid the need for acidic deprotections which are required by some other protecting groups, such as *tert*-butoxycarbonyl (BOC).



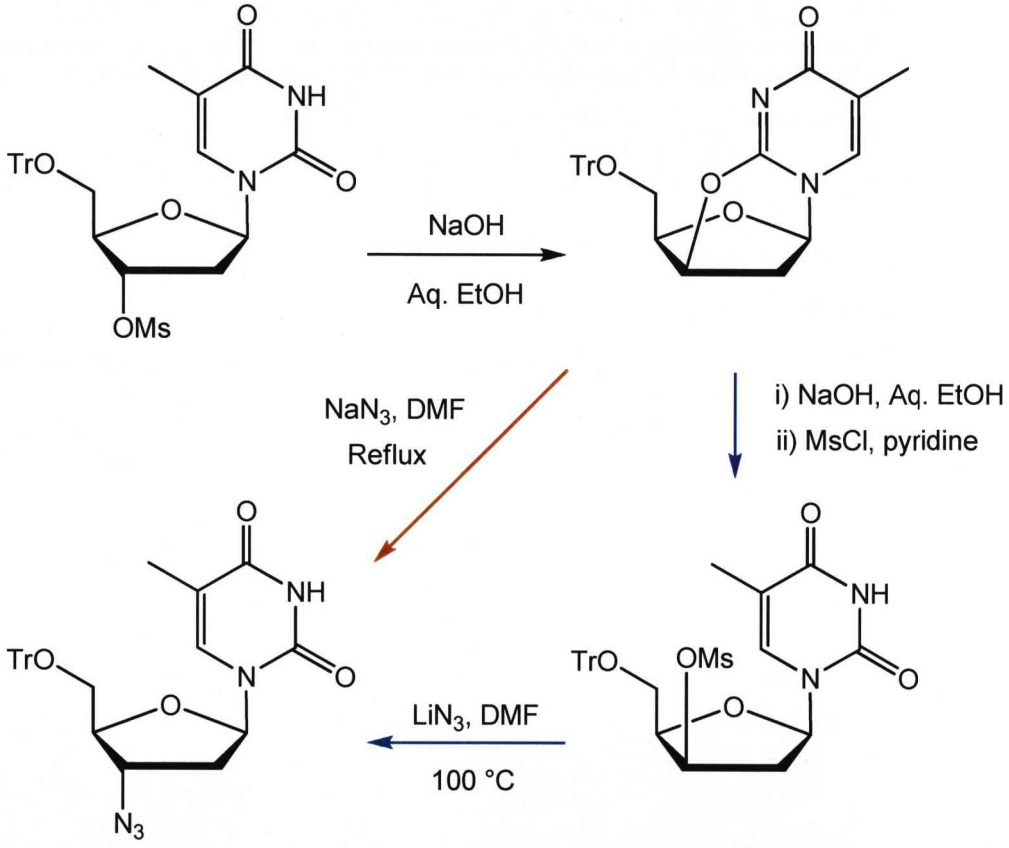
1 **2** T = Thymin-1-yl

Figure 40

Having been so intensively studied as the frontline defence against HIV,[162] a variety of syntheses of AZT were available in the literature.[163] The range includes an 11 step total synthesis method beginning from 3'3-diethoxypropionate,[164] carbohydrate-based approaches[165] and modifications of thymidine via 2,3'-anhydronucleosides[166] or 3'-*xylo*-*O*-methanesulfonyl derivatives.[167,168] The latter two methods have proved to be the most popular routes and first appeared in the early reports of AZT in the 1960's & 1970's. The underlying strategy for both of these reactions is essentially the same. Two S_N2 reactions are performed at the 3'-carbon centre by converting the 3'-OH to a 3'-*O*-methanesulfonyl (3'-OMs) derivative by reaction with methanesulfonyl chloride.

2,3'-anhydro compounds were prepared by Glinski and co-workers[166] by displacement of 3'-OMs, using neighbouring group participation from the

heterocyclic base[169] and the anhydro species subsequently opened with lithium azide. The alternative method also generated the 2,3'-anhydro species in the same manner, but instead of direct displacement, both Horowitz[167] and later Lin,[168] chose to open the anhydro and isolate a protected 3'-*xylo*-thymidine before performing a second S_N2 with another equivalent of methanesulfonyl chloride and lithium azide (see *scheme 1*).



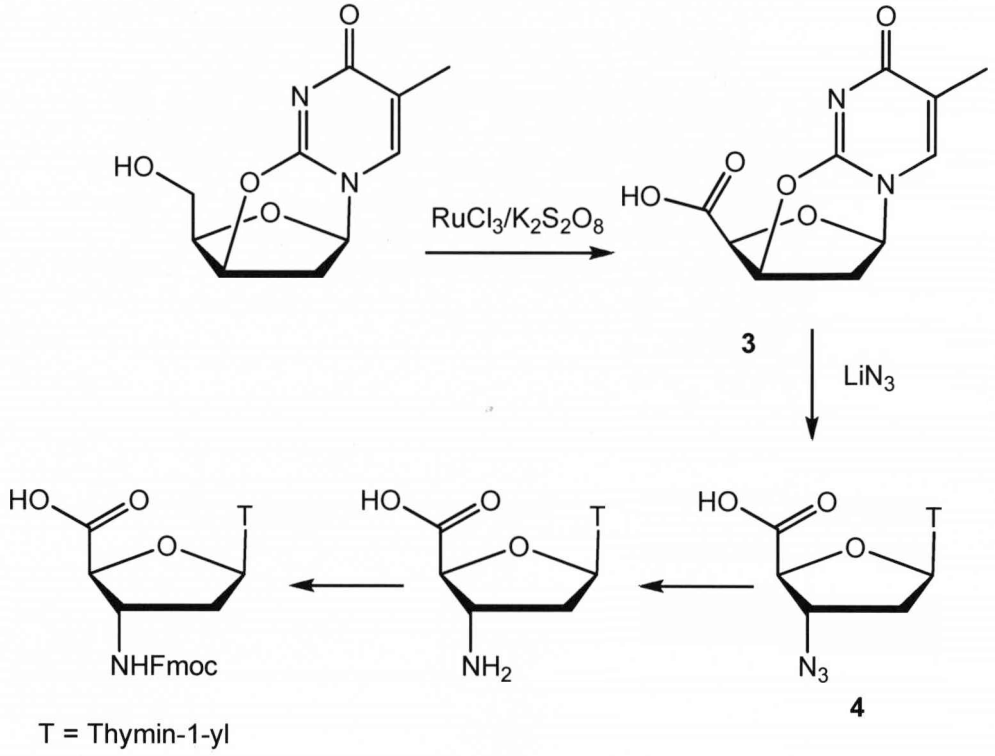
Tr = Trityl or trityl derivative

Scheme 1: *Synthetic routes to AZT by Glinski (red) and Horowitz/Lin (blue)*

At the same time, Fox and co-workers produced 3'-amino-3'-deoxythymidine via 2,3'-anhydro intermediates by reaction with phthalimide.[169] Also, Glinski, Horowitz and Lin all achieved 3'-amino-3'-deoxy compounds by reduction of the 3'-azido nucleosides.

The 2,3'-anhydro compounds have now mostly replaced the dual mesylate S_N2 approach in pyrimidine nucleosides and have since proved to be popular as tools for substitution of 3'-OH with a variety of nucleophiles whilst retaining the natural stereochemistry[170]. The only previous literature syntheses of the target

nucleoside of this project[171] and the related, free amino acid[172] both appear in patent applications and both report azide insertion via a 2,3'-anhydro species.

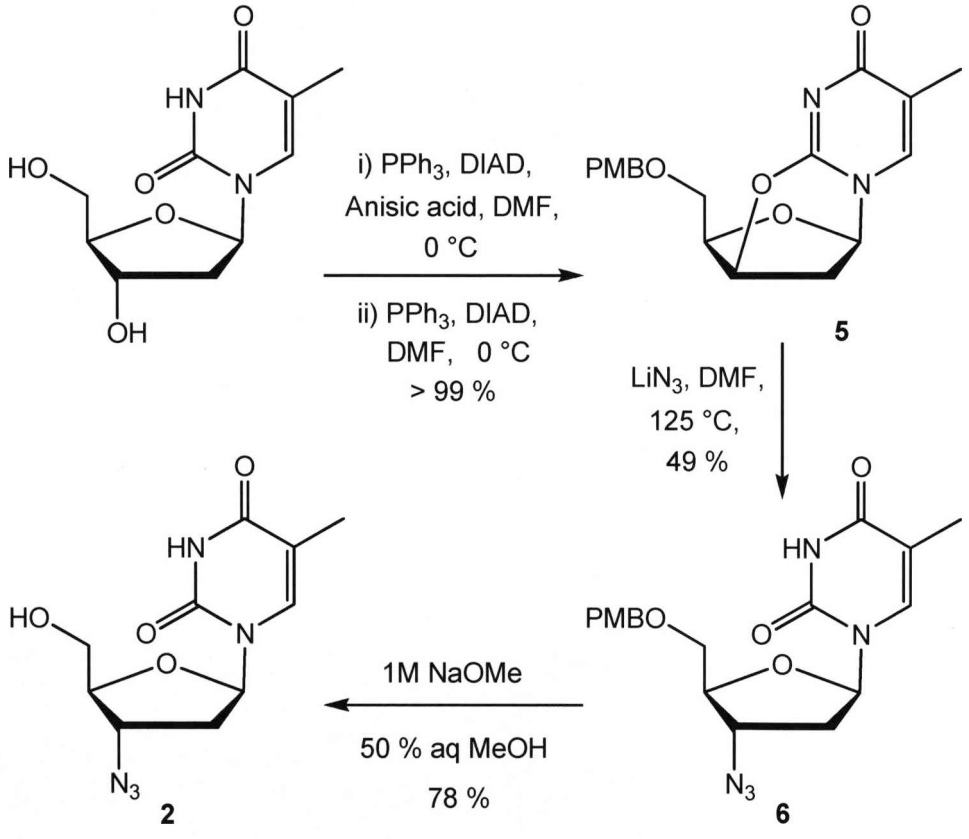


Scheme 2

Varma and co-workers developed two routes to azido acid (**4**) using $RuCl_3/K_2S_2O_8$. The first and simplest was oxidation of commercially available AZT. The same azido acid compound could also be produced by oxidising the 2,3'-anhydro species (**3**) with the $RuCl_3$ system then opening the heterocycle with LiN_3[171] (see *scheme 2*). Cole and co-workers also began with a commercial sample of AZT and oxidised en route to the free amino acid. However, in their work on the 2'-deoxycytidine derived free amino acid, the 2,3'-anyhdro was formed by reaction, not with methanesulfonyl chloride, but with $(PhO)_2SO$. This was then opened from the α-face by nucleophilic attack with azide.[172]

For this study, the method of choice for AZT production was Czernecki and co-workers' elegant tandem Mitsunobu approach[173] (see *scheme 3*). Sequential protection of the 5'-OH with *para*-methoxybenzoic acid then 2,3'-anhydro formation could be achieved by two successive Mitsunobu reactions[174] performed in the same reaction vessel.

1.5 Equivalents of PPh_3 were charged to a vessel containing thymidine and *para*-methoxybenzoic acid in anhydrous DMF solution and excess di-*iso*propylazodicarboxylate (DIAD) added drop wise over approximately 30 minutes. This process was then repeated with the omission of *para*-methoxybenzoic acid. (**5**) Was obtained pure by addition of a large volume of diethyl ether, to force the product to precipitate out of solution, and subsequent filtration. The yield of this reaction was optimised to quantitative by reducing the temperature during DIAD additions to 0 °C.



Scheme 3

The increased yield at low temperature could possibly be ascribed to an increase in the half-life of the reactive nitrogen species generated from DIAD or the stability of the 3'-*O*-phosphonium intermediate with respect to elimination to form the 2'-3'-dideoxy compound during the second Mitsunobu procedure. The unusual ^{1}H NMR shift for the H3' proton of 5.41 ppm was the most indicative sign of successful 2,3'-anhydro formation. Due to extended conjugation through the pyrimidine base the large electron withdrawing effect could also be observed in ^{13}C NMR where C3' saw a downfield shift of 7 ppm compared to thymidine.

Opening of the anhydro species was conducted with excess azide in anhydrous DMF at 125 °C. Lithium azide is generally thought to be superior to sodium azide in this type of displacement due to enhanced solubility in polar solvents such as DMF.[175] In this case the yield with a lithium cation was indeed marginally greater, at around 65 %. The addition of benzoic acid as a proton source to assist in reprotonation of the pyrimidine N3[176] brought no improvement in reaction yield. However, losses of material during a consistently problematic work-up of reactions containing benzoic acid may account for this observation. It has been suggested that improved yields for such anhydro opening reactions may be obtained by alkylating at N3 with a removable group such as benzyl triflate[177] as an alternative to acidic reprotonation. Such methods have yet to prove significantly beneficial.

The reported 90 % yield of protected AZT *via* the anhydro route was not approached over the course of many attempts to produce the 3'-azido compound. Recovery of small quantities of anisic acid during chromatography of the final product suggested that the 5'-OH protecting group was not completely suited to the harsh reaction conditions, despite claims in the original paper that no ester bond cleavage was observed.

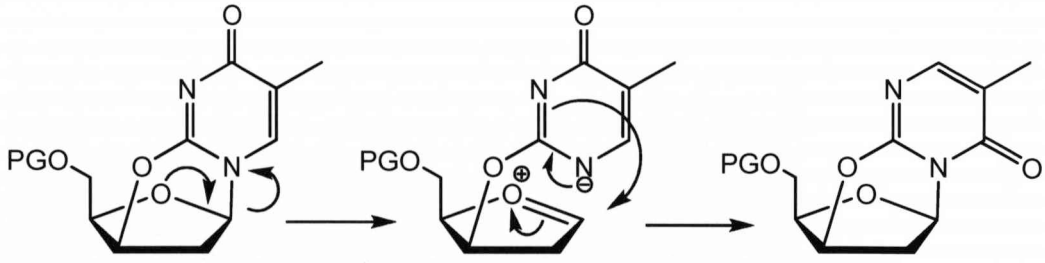
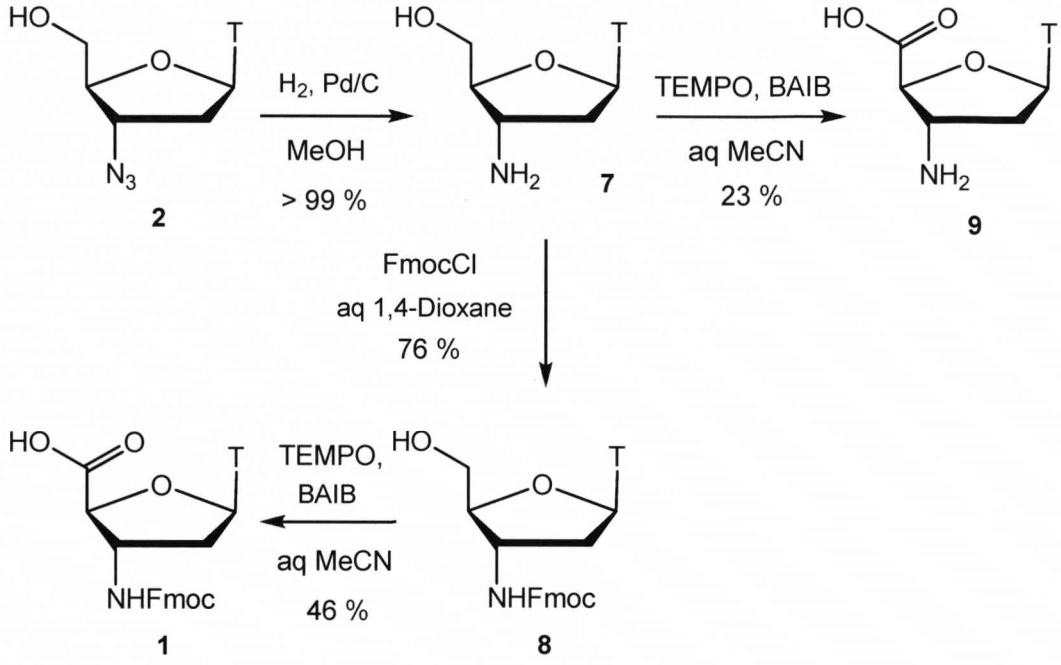


Figure 41: *rearrangement of 2,3'-anhydrothymidine.*

Rearrangement to *N*-3 thymine isomers *via* glycosylic bond cleavage is also possible under the conditions used[178] (see *fig. 41*). No evidence for the formation of such by-products was obtained; although this does not eliminate the possibility of degradation rather than isomerisation resulting from breaking of the glycosylic linkage. The appearance of an azide asymmetric stretch at 2099 cm^{-1} in the IR spectrum confirmed the insertion of N_3. A simple deprotection of the 5'-O-*para*-methoxybenzoyl group with NaOMe in aqueous MeOH followed by neutralisation with Amberlite® H^+ ion exchange resin afforded AZT in 46 % yield from thymidine.

In Cole's methodology[172] the route to the desired β-amino acid from AZT proceeded by oxidation of the 5'-C position with $K_2S_2O_8$ and catalytic $RuCl_3$,[179] then reduction of the azido function by hydrogenation over palladium. Finally, Fmoc protection of the resulting amine with Fmoc chloroformate afforded the target amino acid.



Scheme 4

As an alternative to oxidation of the 5'-carbon followed by reduction of the azide, to enable access to the Fmoc amino acid *via* the free amino acid,[180] it was decided to investigate the possibility of reducing of the azide moiety and protecting the resulting amine prior to 5'-oxidation (see *scheme 4*).

Although alternative reductions for the transformation of azide to amine, such as the Staudinger reaction[181] and H_2S[182] were available, catalytic hydrogenation appeared to be the simplest and most efficient method.[183] It is known that both pyrimidine carbonyl groups and double bonds are stable to hydrogenation using palladium catalysis[184] and this was the natural choice for a highly functionalised nucleoside containing both such groups. Hydrogenation of the azide catalysed by palladium on activated charcoal suspended in MeOH solution, gave 3'-amino-3'-deoxythymidine in quantitative yield within 2 hours. Disappearance of the azide absorption in the IR spectrum coupled with another significant shift of 0.8 ppm upfield for H3' confirmed the formation of the amine (**7**). (**7**) Was then easily protected with 1.3 equivalents of FmocCl in 33 % aqueous 1,4-dioxane to give the

masked amino alcohol (**8**). In the proton NMR, de-shielding by the newly formed carbamate group resulted in a downfield shift of 0.7 ppm for the H3' proton. This resonance was now flanked on either side by signals belonging to the protons bonded to the saturated carbons of the Fmoc group. Appearance of the fluorenyl aromatic peaks distributed between 7 and 8 ppm also confirmed successful protection.

There are a multitude of procedures in the literature for oxidation of the 5'-C position of a nucleoside and traditional transition metal-based procedures remain a common method. The most popular reagent of this type is pyridinium dichromate[185] (PDC). The very similar pyridinium chlorochromate (PCC) does not appear frequently and $KMnO_4$ uses relatively harsh conditions which are not favourable for many nucleoside based transformations.[186] Heavy transition metal-catalysed formation of carbonyl groups is also well tolerated by nucleosides, reliable routes include persulfate with $RuCl_3$[179] and the more direct O_2 insertion over activated platinum in aqueous $NaHCO_3$.[187]

Organic, metal-free oxidation systems are increasing in popularity due to the problems associated with transition metals in terms of purification and over oxidation. Amongst the less mainstream options for nucleosides is biotransformation using the enzyme nucleoside oxidase. First isolated by Isono and co-workers from *Stenotrophomonas maltophilia*[188] the enzyme has subsequently been shown to tolerate a wide variety of synthetic and natural nucleobases with a rapid substrate turnover,[189] but has limited application due to poor stability at room temperature.

As with the nucleoside oxidase reaction, the primary product of most metal-free oxidation reactions performed with nucleosides are aldehydes; an additional oxidant is required to generate the carboxylic acid. A good example of such a reaction was developed by Moffatt and co-workers using a sulfur ylide derived from the reaction of DMSO with dicyclohexylcarbodiimide (DCC) in the presence of pyridinium trifluoroacetate.[190] In the original paper the aldehyde produced was oxidised further to the carboxylic acid for characterisation purposes with periodate. This can also be achieved with other co-oxidants such as mCPBA[191] or $NaClO_2$.[192]

The search for even milder conditions for metal-free oxidations has led to the popularity of hypervalent iodine compounds for this purpose because of the ease of access to aldehydes with very little over oxidation, unless co-oxidants are specifically added. The Dess-Martin procedure employs a cyclic periodinane (see *fig 42*) offering short reaction times, easy separation of products and it is selective for

alcohols in the presence of other easily oxidised groups such as furan rings or sulfides.[193] The mechanism of the oxidation is one of ligand exchange at the I centre followed by proton exchange between the substrate and one acetate ligand. The newly oxidised product is then eliminated with corresponding reduction of I(v) to I(iii).

The Dess-Martin reagent is not selective between primary and secondary alcohols and also cannot be applied catalytically. These are the properties which make the combination of another hypervalent iodine complex, *bis*-acetoxyiodobenzene (BAIB) with the free radical 2,2,6,6-tetramethylpiperidinyloxy (TEMPO), such a powerful and attractive system for oxidation.[194] The most attractive aspect of the TEMPO/BAIB system in terms of our target compound was the direct access to carboxylic acids, when the reaction is performed in aqueous acetonitrile, and the known propensity of nucleosides to precipitate out of the reaction solution without requiring further purification.[195]

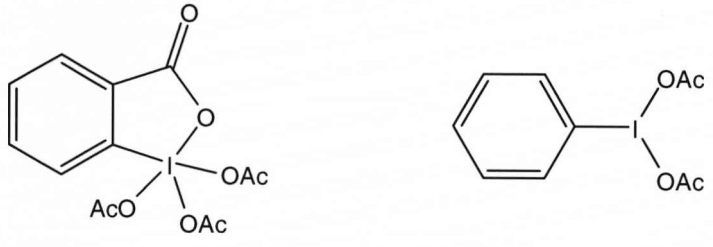


Figure 42: *Dess Martin periodinane (left) and bis-acetoxyiodobenzene (right).*

Nitroxyl radicals such as TEMPO have been identified as useful reagents for oxidations using a number of different conditions and co-reagents over a remarkably large range of substrates.[196] As expected of a metal-free technique, the product of TEMPO oxidations is an aldehyde with additional reagents required to oxidise further to an acid. TEMPO can be applied in oxidation reactions in two ways. It has been identified that the active species for oxidation is not the free radical itself, but the oxoammonium species[197] and thus it is possible to consider utilising either a stoichiometric or a catalytic quantity of TEMPO. In a stoichiometric reaction the oxoammonium salt can be either prepared and isolated separately from the substrate[198] or generated *in situ* by use of an additional oxidant such as mCPBA.[199] Catalytically applied, only a small amount of the free radical is required and this can be converted *in situ* to the oxoammonium species by a co-oxidant in a catalytic cycle. It is these systems that are the most mechanistically interesting that have been the most intensively studied.

The absolute mechanism of the overall conversion remains uncertain, however, there are two definite pathways for the generation of the oxoammonium species; *syn* and dis-proportionation.[200] Below pH 2, disproportionation is favoured. This is induced by the acidic conditions and the radical is regenerated by the co-oxidant from the hydroxylamine (see *fig. 43*).

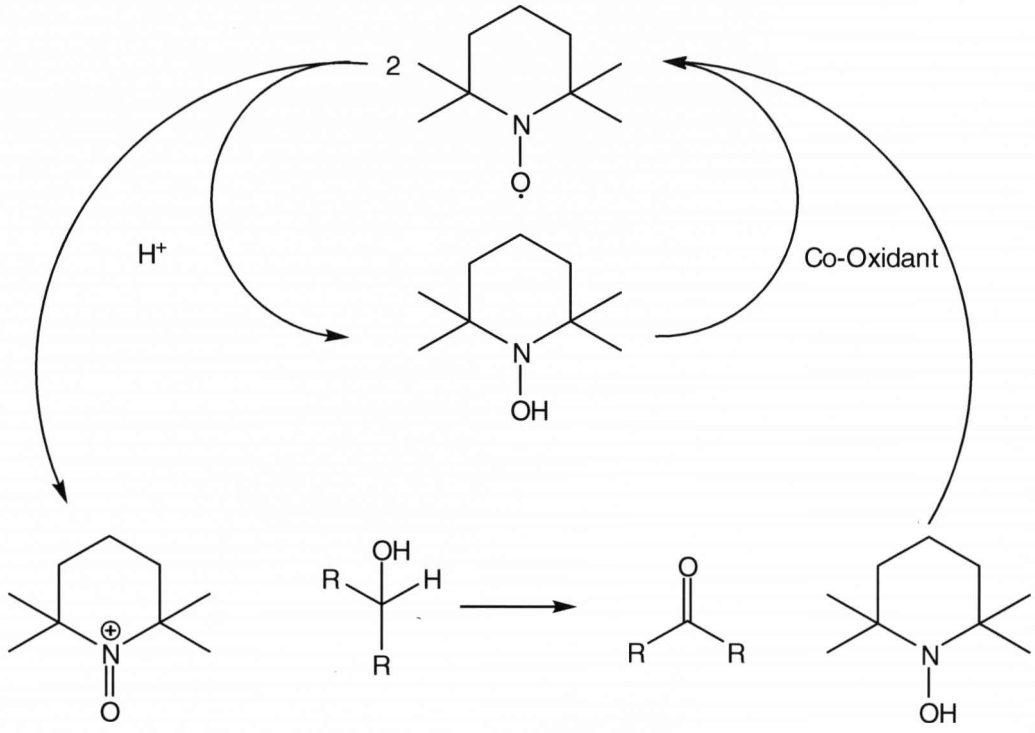


Figure 43: *Disproportionation*

Above pH 3, it is suggested that *syn*-proportionation of the oxoammonium salt and hydroxylamine occurs to give two nitroxyl radicals which are then oxidised by the co-oxidant to the active oxoammonium salt (see *fig. 44*). A variety of co-oxidants can be used in both *syn*- and disproportionation cycles including $NaOCl/Br_2$,[201] $CuCl/O_2$,[202] Oxone[203] and oxidation can also be achieved electrochemically.[204] In the case of BAIB it is thought that, even though no extra acid is used, the disproportionation route is almost certainly followed due to the generation of acetic acid upon ligand exchange with the substrate at the I (iii) centre. In this reaction BAIB has a dual role as both the co-oxidant and source of protons. The nitroxyl radical is then converted to the oxoammonium salt and the aldehyde is formed by one of three possible intermediates (see *fig. 45*).

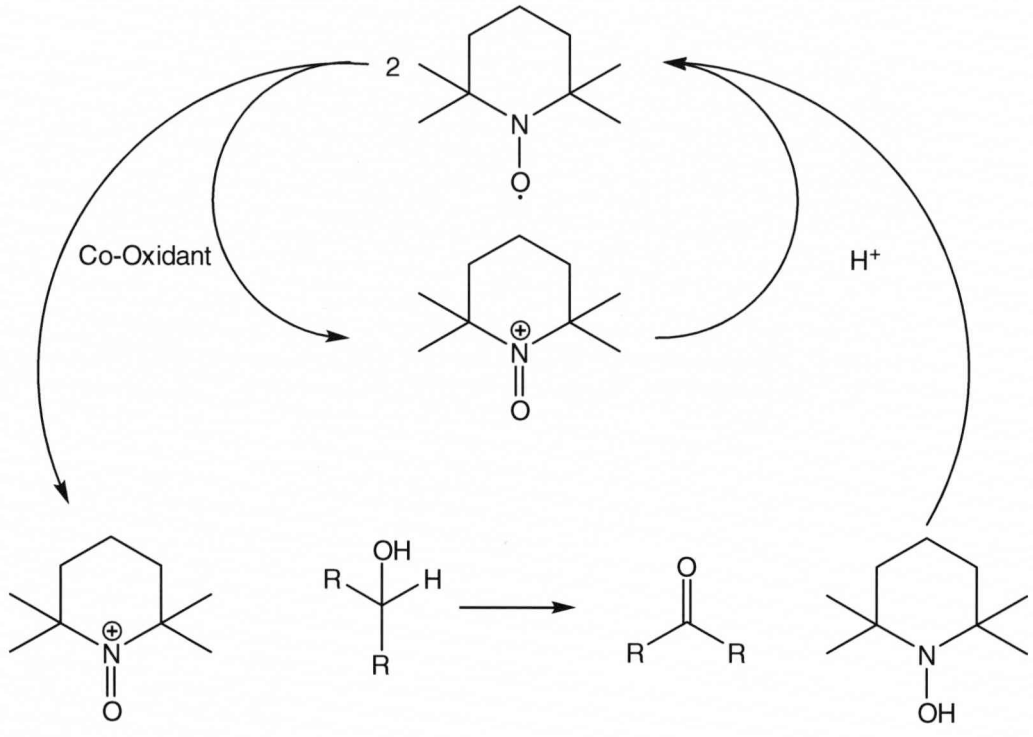


Figure 44: *Syn-proportionation*

When the Fmoc amino nucleoside was treated with TEMPO/BAIB in 50 % aqueous acetonitrile the corresponding carboxylic acid (**1**, p.48) did indeed precipitate from the reaction solution and was obtained in 40 – 50 % yield, without the need for purification. Disappearance of the H5' protons from the ^{1}H NMR coupled with the shift of C5' to 172 ppm made for easy characterisation of the protected amino acid. The free amino acid (**9**, *scheme 4*) was also produced in 23 % yield by applying the TEMPO oxidation conditions to amino thymidine (**7**).

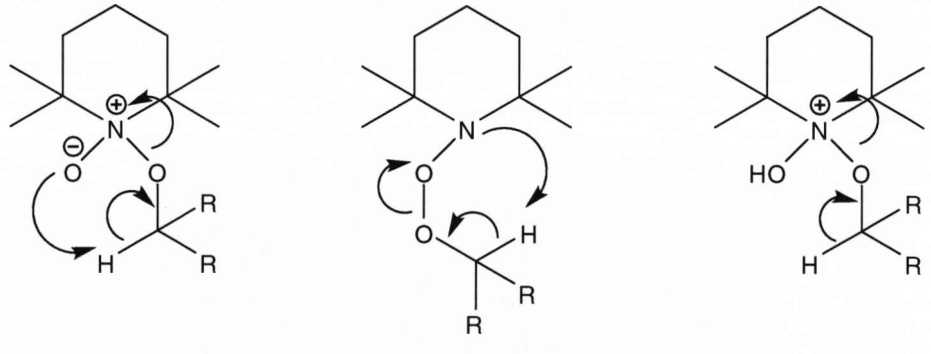


Figure 45: *Possible reaction mechanisms of TEMPO oxidation.*

Attempts to determine the fate of the remaining 50 % of the starting material proved unsuccessful. The only recognisable products in the reaction liquors being

the expected by-products acetic acid and iodobenzene. A possible explanation for the loss of the starting material came to light when Miyamoto and co-workers identified hypervalent iodine compounds, including BAIB, as an ozone equivalent for oxidative cleavage of carbon – carbon double bond in the presence of water.[205] The double bond of thymine is known to be susceptible to ozonolysis[206] and it is plausible that exposure to BAIB in water led to decomposition of the starting material. However, no physical evidence has been obtained to support this theory.

2.2 Solid-Phase Peptide Synthesis of Thymidine-derived β-Peptide

As alluded to earlier, the Fmoc group was chosen for amine protection in order to conserve the nucleoside glycosylic bond. Acidic conditions could not, however, be completely avoided but resin and reagents for SPPS were selected with minimum acidic conditions in mind. The resin chosen for solid phase synthesis was Sieber amide resin with the hyper-acid labile xanthen-9-amine linker group[207] (see *fig. 46*). This enabled cleavage of the oligomer from the solid support to be achieved under very mildly acidic conditions, 1 % trifluoroacetic acid (TFA) in DCM solution, minimising the chance of cleavage of the glycosylic bond. The Sieber amide resin, as the name suggests, produces terminal amides upon cleavage as opposed to carboxylic acids. For preliminary synthetic studies this was not perceived to be of any significant consequence and SPPS was thus pursued with this resin.

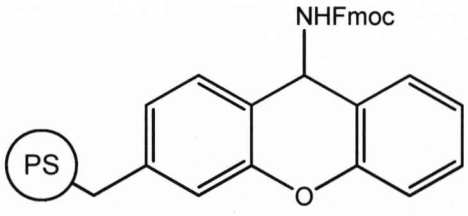


Figure 46: *Sieber Amide Resin.*

A common problem with β-peptides and PNA oligomers alike is their poor solubility in aqueous solvents due to the neutral nature of the peptide backbone[90,125]. Initially it was believed that oligomers derived from the thymidine β-amino acid could also be insufficiently soluble in aqueous solution to allow either purification by reverse-phase HPLC or subsequent structural studies to be performed. A common strategy to increase solubility of PNA type peptide chains is to include additional lysine residues within the chain.[118,127] Tagging either the *C* or *N* terminus of PNA oligomers with Fmoc-protected hydrophilic natural α-amino acids is known not to

known not to hinder hybridisation with complementary DNA strands.[125] For our purposes it was most appropriate to include any additional residues at either terminus so as to minimise any impact on secondary structure formation. It was therefore decided to append each end of the peptide chain with a natural *L*-lysine residue to maximise the solubilising effect. The initial SPPS experiments with the thymidine derived amino acid were conducted with this in mind (see *fig. 47*). During the peptide synthesis itself, this amine side chain would require protection to prevent unwanted side reactions. An orthogonal, acid-labile monomethyltrityl (Mtt)[208] group was chosen for this purpose. This enabled both the peptide and the side chain protecting group to be cleaved from the acid sensitive resin concurrently at the end of the synthesis to give a fully deprotected oligomer chain.[209]

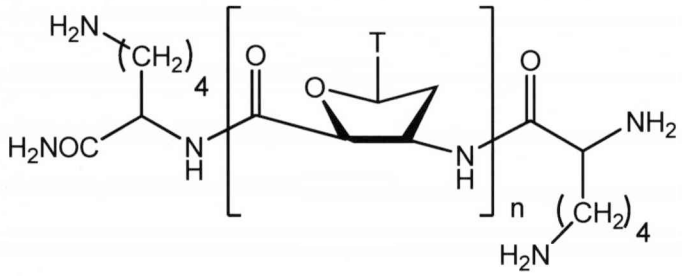


Figure 47: *Nucleoside β-peptide appended with L-lysine residues.*

Peptide synthesis was conducted as follows: The resin was swelled overnight in DCM then drained and washed with DMF. The Fmoc protecting group on the resin was then removed by agitating the resin with 20 % piperidine in DMF solution before again draining and washing the resin bed. The amino acid to be coupled was then dissolved in DMF and pre-activated for 60 seconds with the amide bond forming reagent *O*-benzotriazole-*N*,*N*,*N*',*N*'-tetramethyluronium hexafluorophosphate (HBTU) and di*iso*propylethylamine before being added to the resin bed and agitated. Upon completion of the coupling reaction the amino acid solution was removed. After washing, the resin bed covered with a "capping" solution containing 10 % acetic anhydride and 10 % 2,4,6-collidine in DMF to acetylate any amines that had not reacted before the next coupling cycle was commenced. At the end of the synthesis the resin bed was washed with DMF, DCM and methanol then dried under vacuum overnight. Re-swelling with DCM for 30 minutes was followed by treatment with 1 % TFA in DCM, to cleave the peptide from the resin, and 5 % triethylsilane (TES) in DCM, to scavenge the monomethyltrityl group. These liquors were drained into a flask containing 10 % pyridine in DCM in

order to neutralise the TFA and, after removal of all the solvents under reduced pressure, the peptide was precipitated by the addition of diethyl ether.

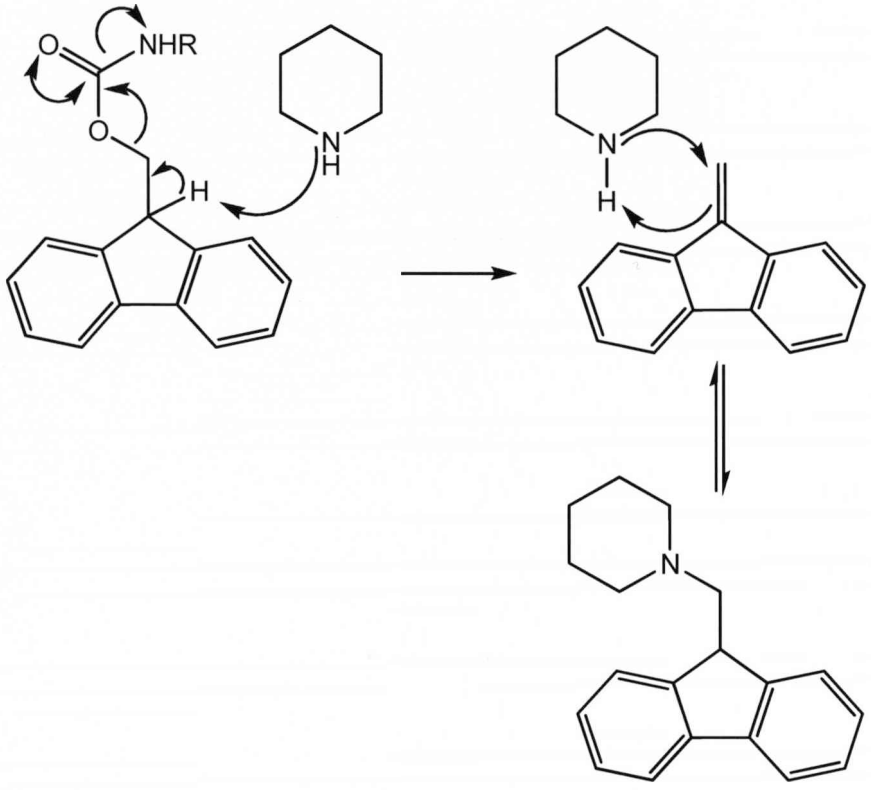


Figure 48: *Mechanism of Fmoc removal and adduct formation with piperidine.*

Initial attempts to construct oligomers of the form Lys-5'-mT_n-3'-Lys using the Sieber amide resin failed to produce a significant amount of peptide, as determined by HPLC of the crude cleavage mixture. It was clear that the coupling conditions would require optimisation. During the optimisation process, coupling yields could be monitored by a UV spectroscopic assay of deprotection solutions for the presence of dibenzofulvene (DBF).[210] DBF is liberated upon exposure of the Fmoc group to base, however, piperidine can add to DBF in a Michael fashion leading to a distortion of the concentration of DBF and, consequently, underestimation of coupling efficiencies[211] (see *fig. 48*). At this point 20 % piperidine solution was replaced as the deprotection agent by a 2 % solution of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU).[212] DBU will not add to DBF in the same manner as piperidine and is considered more accurate for these purposes.[213]

As both terminal residues of the peptide were to be lysine, the first coupling reaction to be optimised was that of Fmoc/Mtt protected lysine with the resin itself. Both HBTU and benzotriazol-1-yl-oxytripyrrolidinophosphonium

hexafluorophosphate (PyBOP) were tested as amide bond forming reagents (see *fig. 49*). Several solvents,[214] DMF, dimethylacetamide (DMAc), *N*-methylpyrrolidone (NMP) and 20 % DMSO in DMF, were also examined for this reaction.

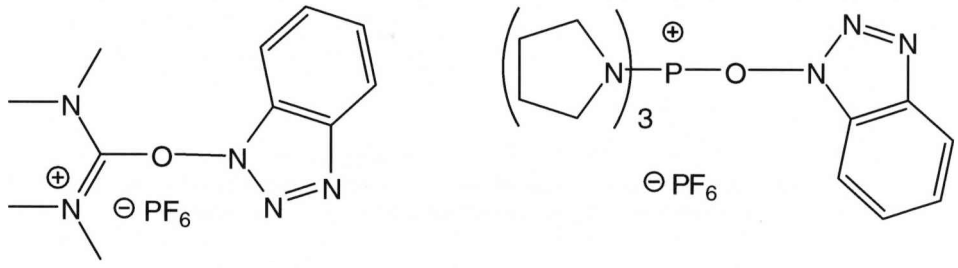


Figure 49: *HBTU (left) and PyBOP (right).*

Table 1 shows the most efficient conditions for the coupling of lysine to the resin were using either DMF or DMAc with HBTU as coupling reagent. HBTU has been shown to give excellent coupling performances with model PNA sequences,[122] however, PyBOP was only slightly inferior as a coupling reagent, as was 20 % DMSO/DMF as a solvent. NMP was a poor solvent in this test. The two best performing combinations were progressed to couplings with the modified nucleoside. The results show that coupling efficiency between the modified nucleoside and lysine was only slightly better with DMF and further improved by the use of PyBOP in place of HBTU. The reason for the failure of the first SPPS attempt using DMF then became clear when the coupling between the modified nucleosides was examined.

Very poor coupling yields were obtained when DMF was employed in combination with HBTU. Performing the coupling reaction twice before the capping step, known as double coupling, also failed to improve yields. PyBOP continued to produce superior yields over HBTU in DMF; however, couplings in DMAc between the nucleosides were quantitative with HBTU. Thus DMAc was introduced in place of DMF. HBTU was retained as the coupling agent as the overall yields for the whole synthesis were projected to be better with HBTU than with PyBOP. The reason for the vast difference between these comparable solvents is unclear. Efficient peptide synthesis is often achieved using these types of polar, aprotic solvents in order to suppress aggregation between chains[215] and retain good resin solvation.[216] The physical and electronic properties of DMF and DMAc are approximately equal and consequently there are no obvious variations in solvating power or electronic effects that could be identified as a possibly cause of this disparity.

Amino Acid[a]	Solvent	Coupling Agent	Coupling Yield / %
Lysine	DMF	HBTU	94
mT1	DMF	HBTU	88
mT2	DMF	HBTU	23
Lysine	DMF	PyBOP	91
mT1	DMF	PyBOP	99
mT2	DMF	PyBOP	92
Lysine	DMAc	HBTU	94
mT1	DMAc	HBTU	86
mT2	DMAc	HBTU	99
Lysine	20 % DMSO/DMF	HBTU	92
Lysine	NMP	HBTU	54

Table 1 [a] The entry "Lysine" refers to the coupling of Fmoc-Lysine-(Mtt)-OH with the resin. "mT1" refers to the first thymidine derived amino acid coupling to lysine and "mT2" to the second thymidine derived amino acid coupling to mT1.

Concurrently with development of suitable reaction conditions, a suitable method of testing the resin at various stages of the peptide synthesis cycle was identified. Solid-phase resin testing during peptide synthesis is commonly performed to test for the presence of free amine groups on the resin. Free amine groups are indicative of a successful deprotection reaction or a failed coupling reaction and can be a useful tool in determining the progress of a given synthesis. Traditionally a Kaiser test[217] is used to show the presence of amines in solid-phase synthesis. It is performed by heating a mixture of 5 % w/v ninhydrin in ethanol solution, 80 % v/v phenol in ethanol solution and 2 % v/v (2 mM KCN solution in water) in pyridine solution at 120 °C for 4 minutes in an ignition tube containing a few of the resin beads. If free amine groups are present, *i.e.* if a deprotection succeeded or a coupling failed, then the beads become stained blue as a ninhydrin adduct forms at the free amine site and subsequent hydrolysis leads to the formation of the compound known as "Ruhemann's purple" (see *fig. 50*). If the beads remain colourless then this indicates that there are no free amine sites present after a failed deprotection or successful coupling reaction. The Kaiser test is, however, not suitable with serine, asparagine, aspartic acid and secondary amino acids such as proline.[218] This was also found to be the case with the modified thymidine

nucleoside. After the first thymidine residue was coupled the Kaiser test invariably gave a pale blue/purple solution through which the colour of the beads could not be determined.

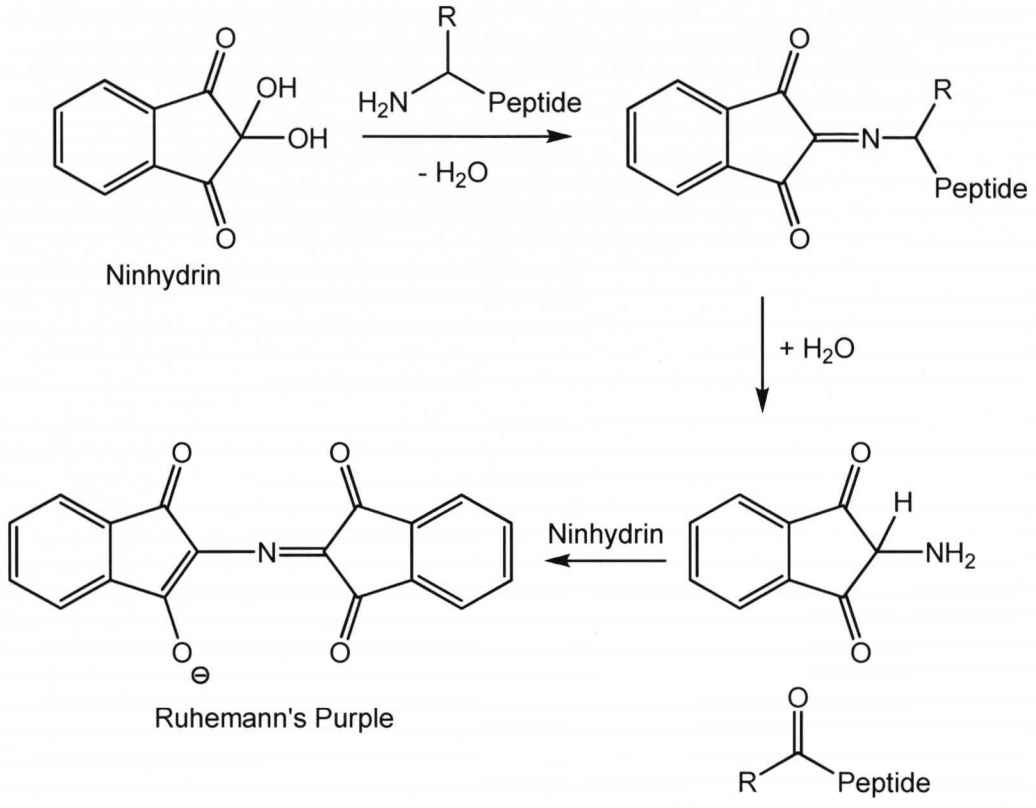


Figure 50: *Ruhemann's Purple from ninhydrin in the Kaiser test.*

Trinitrobenzene sulfonic acid (TNBS)[219] and chloranil (see *fig. 51*) are both well known alternative reagents for resin testing of this type, chloranil being especially noted for use with proline [220] Both of these tests were initially superior to the Kaiser method in that visual contrast between test solution (bright yellow) and beads testing positive for free amines (blue for chloranil, red for TNBS) was high. However, chloranil could not detect free amine sites on the thymidine residues leaving TNBS, despite the handling hazards, as the most suitable resin testing reagent.

With a now optimised coupling process and reliable methods for both monitoring coupling efficiency and amine visualisation, further SPPS to produce lysine appended peptides was attempted. DBF assay during the synthesis showed that, although coupling yields were lower than may have been expected, the 6-

residue peptide was prepared in 15 % overall yield. UV analysis of the post cleavage mixture following the synthesis indicated the presence of thymine with a significant absorption at 256 nm. HPLC analysis showed a single large peak with a retention time of 14.4 minutes. However, when subjected to electrospray mass spectrometry, no parent ions for peptides of any length could be identified in either the crude cleavage mixture or samples purified by HPLC. The reason for the failure of these peptides to fly during mass spectrometry is not known. One hypothesis is that the presence of residual compounds or salts from the cleavage procedure or HPLC purification overwhelmed the relatively weaker, multi-charged ions of peptide spectra. In an attempt to circumvent this characterisation problem two alterations were made to the SPPS system.

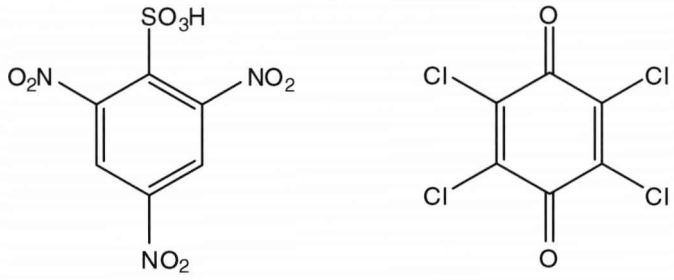
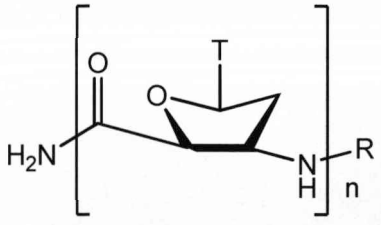


Figure 51: *Trinitrobenzenesulfonic acid (left) and chloranil (right).*

Firstly, the cleavage procedure was revised by removing the TFA by rapid evaporation under reduced pressure instead of neutralising with pyridine, in order to eliminate the presence of salts. The second change was to discontinue the use of lysine as the terminal residues with the aim of reducing superfluous compounds in the final cleavage mixture *i.e.* Mtt derived TES. This also decreased the probability of producing failure sequences, diminishing the need for purification by HPLC.



(**10**) n = 2, R = H
(**11**) n = 4, R = H
(**12**) n = 4, R = Fmoc
(**13**) n = 8, R = H
(**14**) n = 8, R = Fmoc

Figure 52: *Nucleoside β-peptides derived from thymidine produced by SPPS.*

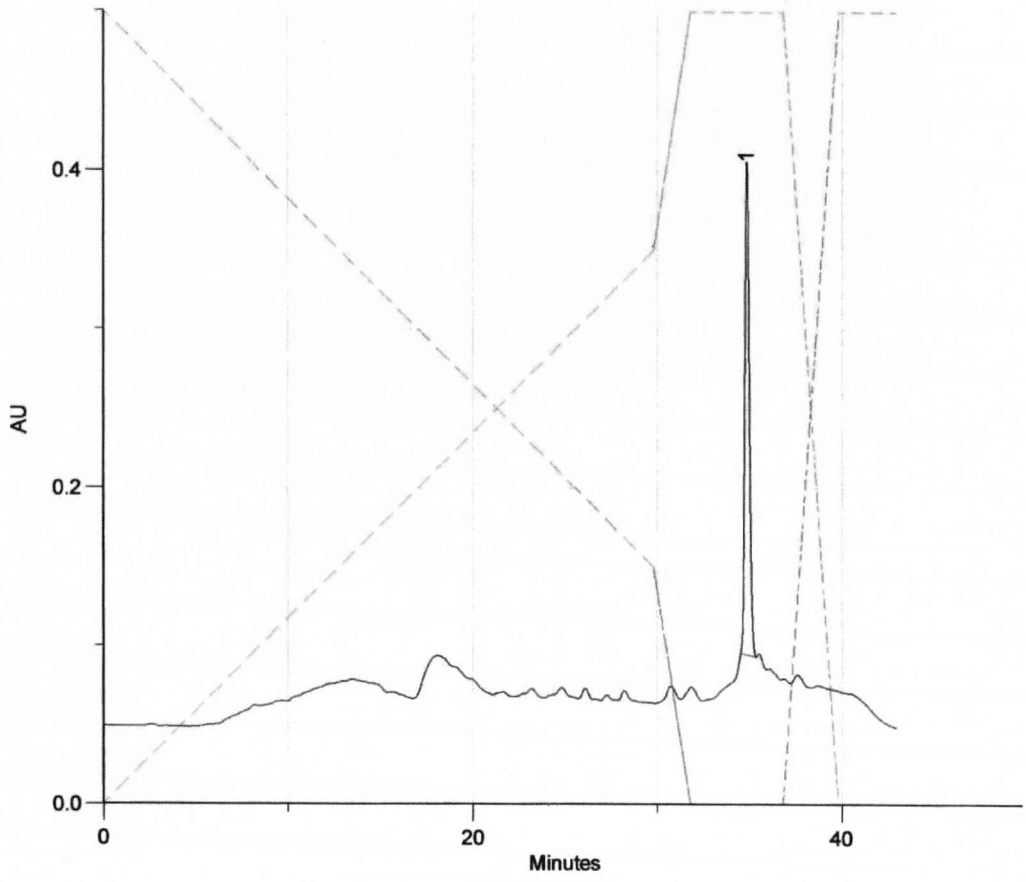


Figure 53: *HPLC trace of a crude sample of peptide (**14**).*

These modifications to the procedure proved successful to the point of obtaining a series of peptide oligomers of the form shown in *figure 52* which were isolated from cleavage mixtures by trituration with diethyl ether. Peptides produced using this modified method were sufficiently pure upon isolation to be analysed by NMR without the need for further purification by HPLC (see *fig. 53*).

Oligomer	n	R	Molecular Formula	HR mass required	HR mass found
10	2	H	$C_{20}H_{26}N_7O_8^+$	492.18	492.18
11	4	H	$C_{40}H_{48}N_{13}O_{16}Na^+$	988.32	988.13
12	4	Fmoc	$C_{55}H_{57}N_{13}O_{18}Na^+$	1210.38	1210.38
13	8	H	$C_{80}H_{91}N_{25}O_{32}Na^+$	1936.62	1936.61
14	8	Fmoc	$C_{95}H_{101}N_{25}O_{34}Na^+$	2158.68	2158.69

Table 2

Peptides **10** - **14** were successfully characterised by high resolution mass spectrometry (see *table 2 and fig. 54*).

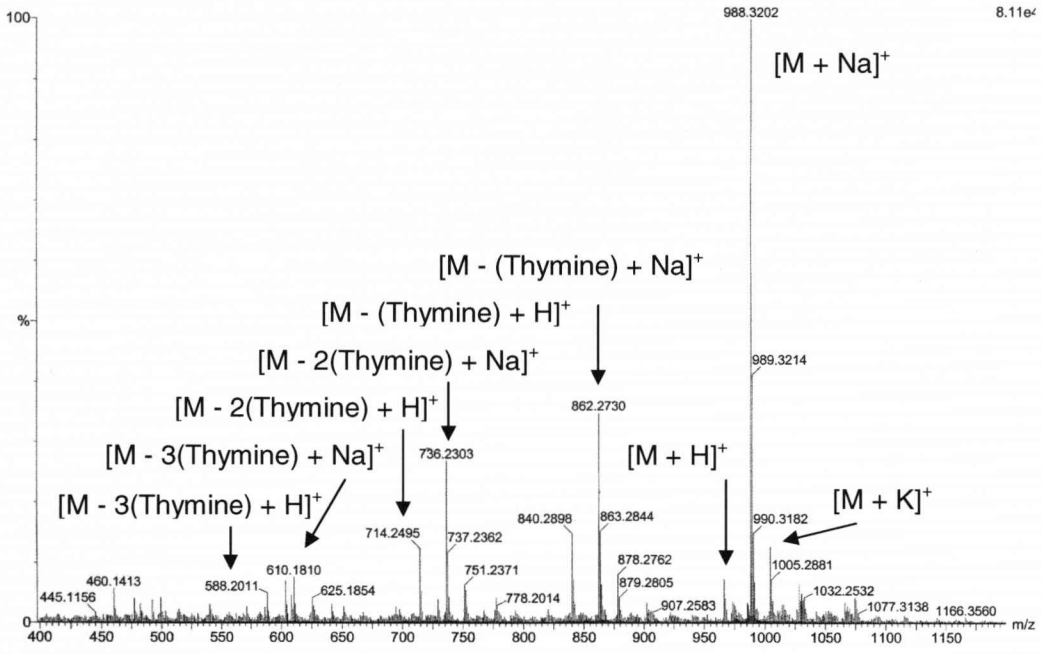


Figure 54: *Electrospray mass spectrum of peptide (**11**)*

With a now reliable method for SPPS and peptide characterisation, the opportunity was taken to examine the viability of alternative resins for conducting SPPS. This was undertaken with the aim of producing C-terminal carboxylic acids instead of amides for possible application in the study of formation of cyclic structures by this class of peptide. Again, taking into account the desire to avoid resins requiring strong acid to cleave the finished peptide, Rink acid[221] and 4-hydroxymethyl-3-methoxyphenoxybutyric acid benzhydrylamine (HMPB-BHA)[222] resins (see *fig. 54*) were tested. Neither of these resins gave superior results when compared to the Sieber amide resin (see *table 3*).

Resin	Solvent	Coupling Agent	Coupling Yield/%
Rink Acid	DMAc	HBTU	10
Rink Acid	DMAc	PyBOP	9
HMPB-BHA	DMAc	HBTU	51
HMPB-BHA	DMAc	PyBOP	60

Table 3

This may be due to changing the connection between the linker and the amino acid from an amide bond to an ester bond. In later, related work on fluorous supports (see chapter 3) it was indeed found that the nucleoside was reluctant to esterify and the choice of the amide resin was, therefore, justified.

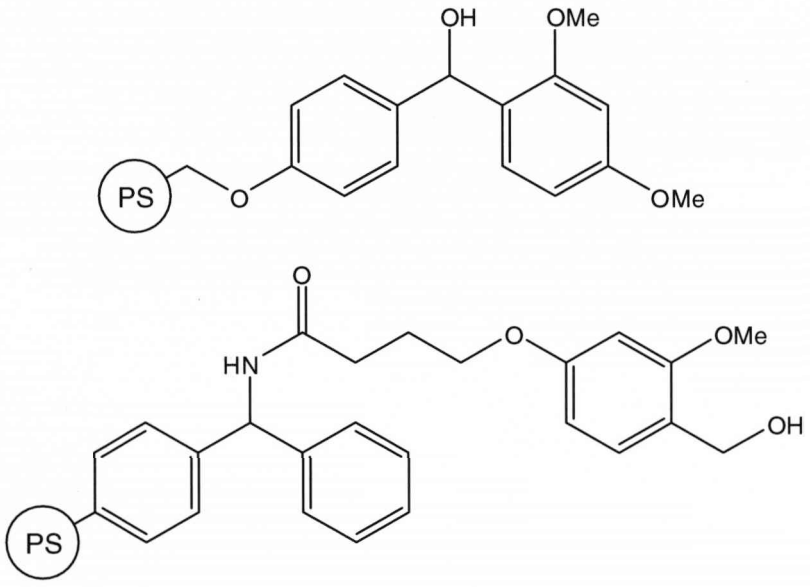


Figure 54: *Rink acid resin (top) and HMPB-BHA resin (bottom).*

2.3 Structural Characterisation of Thymidine-derived β-Peptide

In collaboration with Leeds University, DFQ-COSY, NOSEY, TOCSY, VT-NMR and UV melting experiments were used to examine the solution conformations of the peptide chains.[223] It was possible to obtain a complete resonance assignment for oligomer (**12**) and, by analogy, assign conformational features of oligomer (**13**). The NH protons of tetramer (**12**) were all unambiguously assigned and it was observed that the three inter-nucleoside amide protons showed a marked difference in chemical shift, 8.85, 8.92 and 8.94 ppm, to that of the terminal, Fmoc protected amine, 8.13 ppm (see *fig. 55*). This was the first indication of an intramolecular hydrogen bonding network incorporating these inter-residue amide protons.[95] A variety of nOe connections were also identified (see *fig. 55*) and these constraints were then included in the construction of a model of the short peptide strand.

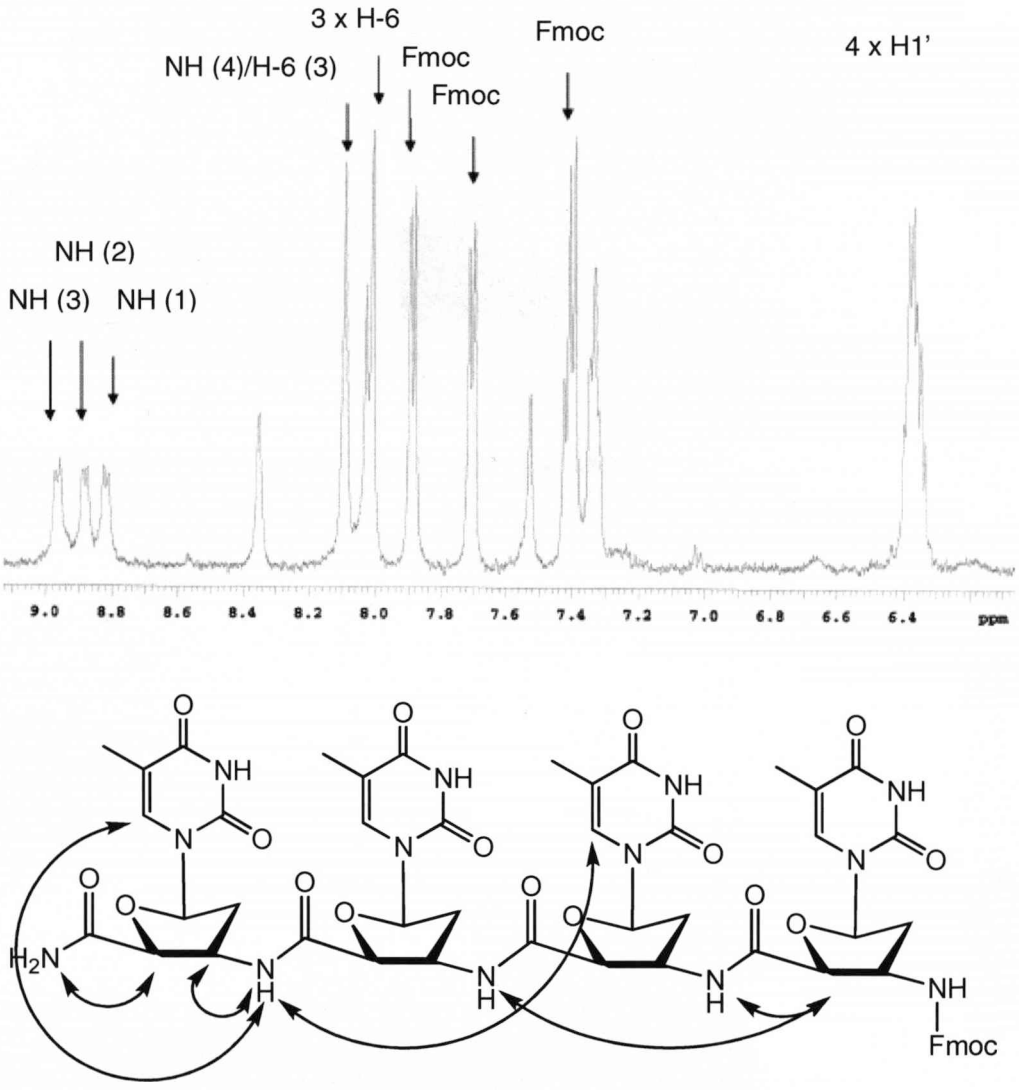


Figure 55: *Aromatic region of 1D NMR spectrum and selected nOe connections observed in peptide (**12**).*

The model suggested that a helix-like pattern was indeed present (see *figs. 56 and 57*). The core of the helix was occupied by a zigzag arrangement of the sugar rings, forcing the sterically demanding thymine bases to protrude into two well-defined planes either side of the chain in a right-handed helical arrangement.

Further evidence of the presence of a hydrogen-bonded network within the peptide was obtained by the observation of larger, high-field shifts of the inter-residue amide protons, approximately 6 ppbK^{-1}, compared to that of the terminal amine, 3 ppbK^{-1}, in variable temperature experiments. This result was to be expected given that the model suggested a near-linear configuration.[224] Combining all these findings

together, examination of the environment of the amide protons showed hydrogen bonding between the amide protons and the carbonyl oxygen belonging to the n+2 residue of the chain. An atom count along the chain revealed an 8-helix.

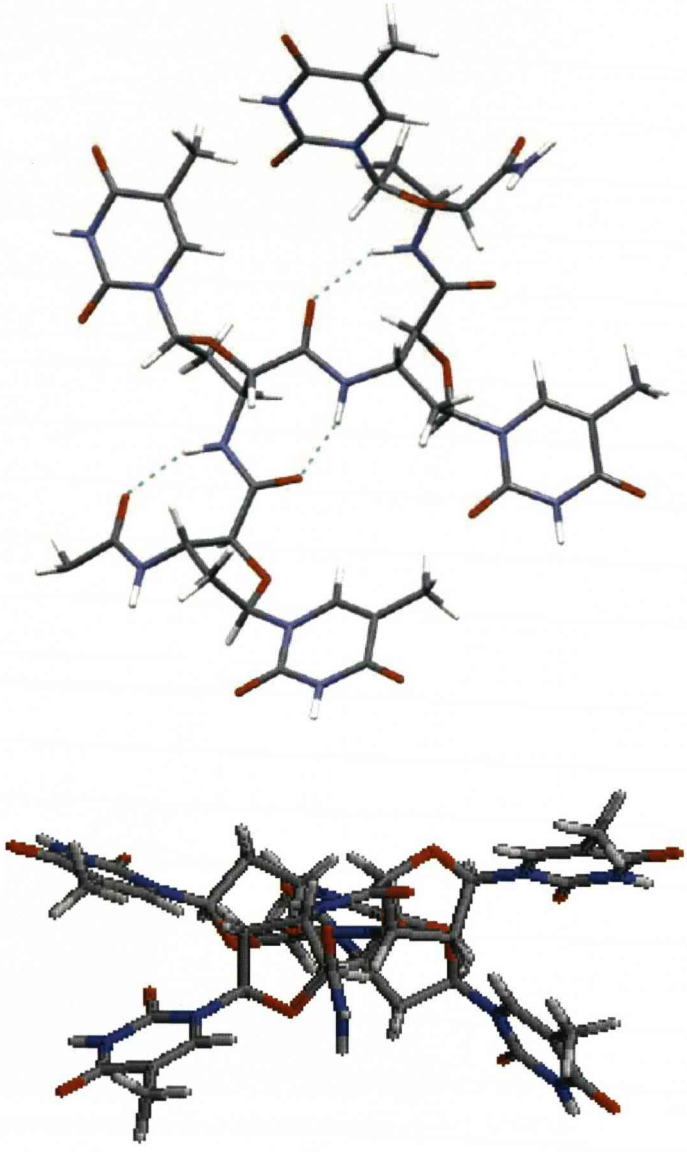


Figure 56: *Side and top view of model of 8-helix formed by peptide (**12**) (model constructed using MacroModel,[225] courtesy of C. Kilner).*

8-Helices, whilst more unusual than 12- or 14-helices, have previously been observed in conformationally constrained cyclic β-peptides.[94,95] It is believed that the constraints in this case are derived from the steric demands of the heterocyclic base.

Similar observations were made when examining octamer (**13**) indicating comparable structural behaviour proliferates within longer peptides. However, with a

much more complicated spectrum, only the resonances belonging to the terminal residues have so far been assigned.

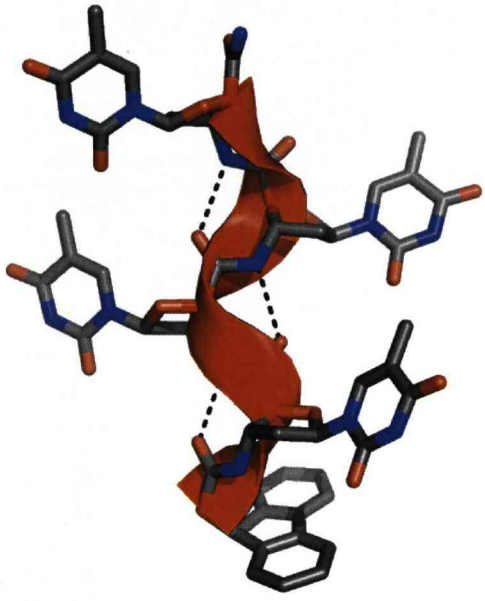


Figure 57: *Alternative model showing helical core of peptide (**12**) (model constructed using Spartan, courtesy of N. Berry).*

The suggestion that ordered structure was present in the peptide chain was supported by UV melting studies. Single-stranded DNA and RNA have UV absorption maxima at approximately 260 nm with a large extinction coefficient. When two single strands of natural DNA or RNA anneal to form a duplex, a drop in UV absorption of the solution containing the strands of up to 30 % occurs due to the hypochromic effect.[10] This phenomenon is caused by base stacking; the interactions of π-electrons between nearest neighbour heterocyclic bases. Dissociation of the duplex caused by heat can, therefore, be observed by measuring the UV absorbance of a solution of the duplex as a function of temperature. This produces a sigmoidal curve as shown in *Figure 58*. The point at which the duplex is 50 % dissociated, *i.e.* the temperature at which the UV absorbance is half the maximum, is known as the melting temperature (T_m). Changes in T_m are commonly used as a measure of the stabilising or destabilising effect on a duplex of a particular synthetic modification contained within one strand of the duplex or to establish whether a nucleic acid analogue can recognise and bind to a complementary strand of natural DNA.

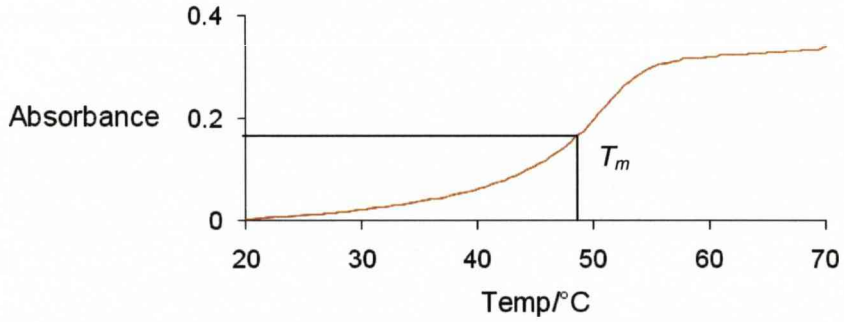


Figure 58: *Example of a melting curve and T_m.*

Melting experiments performed with single strands of (**12**) and (**13**) produced absorbance curves which were similar to those observed for a natural duplex. Although somewhat irregular, the sigmoidal trend of the graphs can be seen suggesting a melting temperature of approximately 51 °C for tetramer (**12** see *fig. 59*) and 57 °C for octamer (**13** see *fig. 60*).

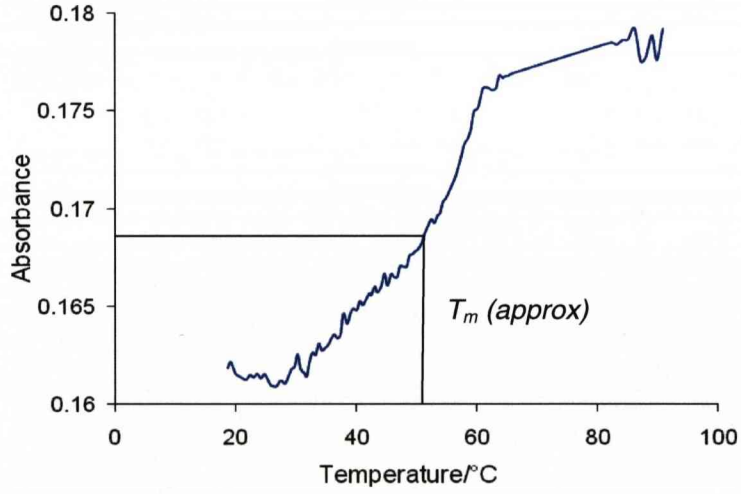


Figure 59: *UV melting curve for peptide (**12**)*

As already discussed, the hypochromic effect in an unmodified duplex is a result of base stacking, however, the model of peptide (**12**) obtained through the NMR experiments suggested that the inter-base distance was between approximately 5.57 – 5.87 Å. This separation is much larger than that observed in either A- or B-form duplexes and is unlikely to support a π-stacking interaction. This implies a more complex intermolecular relationship, such as interdigitation of the bases, as a more likely source of the melting behaviour. More extensive investigation is required to examine this hypothesis.

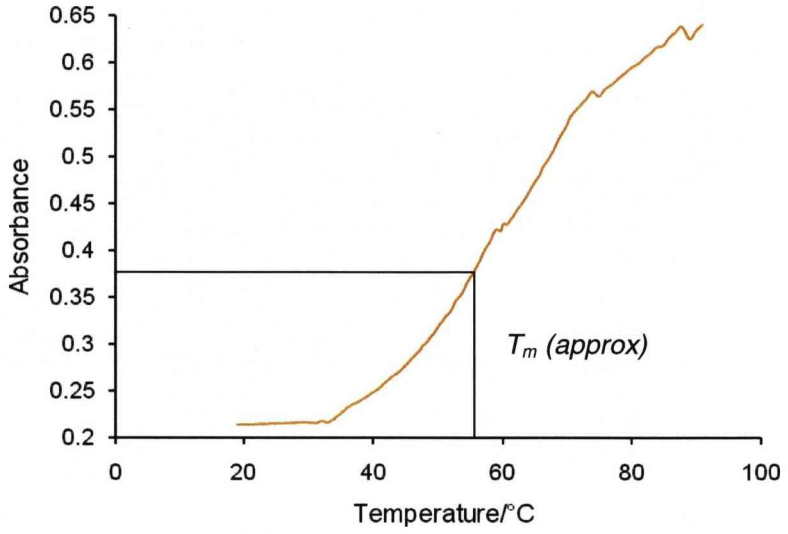


Figure 60: *UV melting curve for peptide (**13**)*

Additionally, a melting experiment with (**13**) in the presence of a complementary strand of homooligomeric dA_8 was also conducted in order to establish the hybridisation properties of the modified strand with natural DNA.

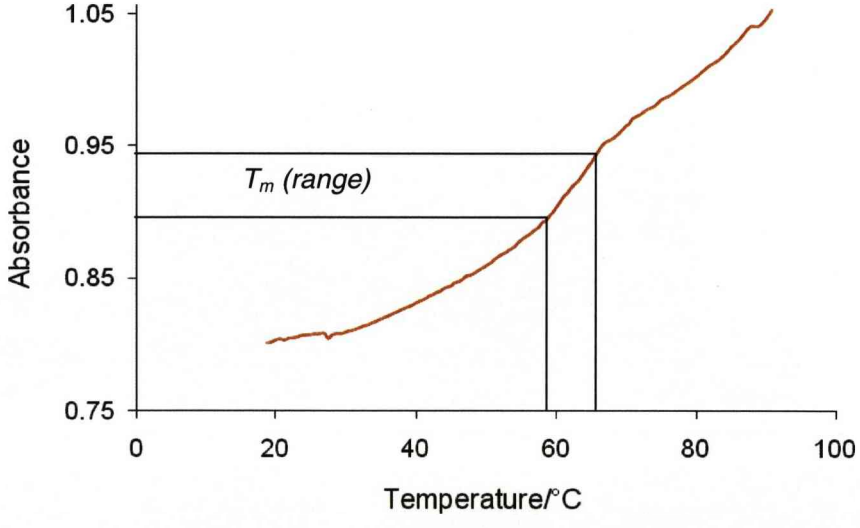


Figure 61: *UV melting curve for modified T_8 peptide strand (**13**) with dA_8 DNA.*

In an unmodified control experiment with T_8/dA_8 DNA strands, a melting curve was not recorded. The absorbance was found to rise linearly with temperature indicating the two strands failed to anneal. This finding was consistent with published data suggesting melting curves for duplexes of this type are usually only observed where high concentrations of salts are present.[226] However, when (**13**)

was subjected to the same experiment with a complementary dA_8 DNA strand, a curve was observed (see *fig.* 61*)*. Again, due to a slightly irregular shape, a melting temperature of between 59 – 64 °C was estimated. It is unclear whether this was due to the denaturation of the structure within the peptide strand, as the value was close to the 57 °C value estimated for the single strand experiment, or to melting of a duplex formed by two strands.

The presence of an 8-helix within the peptide chain was interesting for several reasons. Firstly, stable secondary structure in peptides as short as 4 residues is rare. The presence of a helix in this case suggested strong steric influences on the conformation. Also the configuration of the sugars and bases compared to natural, double-stranded DNA presented the remarkable structural feature of having two planes of exposed heterocyclic bases either side of the helical core. This raised the possibility of conducting hybridisation studies, not only with a strand of complementary 2'-deoxyadenosine (dA) DNA homo-oligomer, but also with a strand of the corresponding dA derived peptide. It could be imagined that, if the dA analogue were to display the same conformational properties as the thymidine derived peptide, the potential existed to create a continuous array of parallel peptide strands held together by Watson-Crick hydrogen bonding. The natural progression of this project, therefore, was towards dA derived peptides produced from the appropriate dA β-amino acid (**15**) (see *fig. 62*).

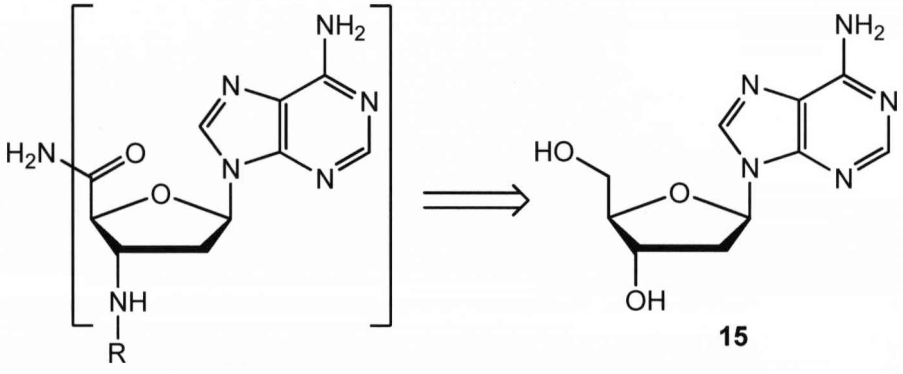
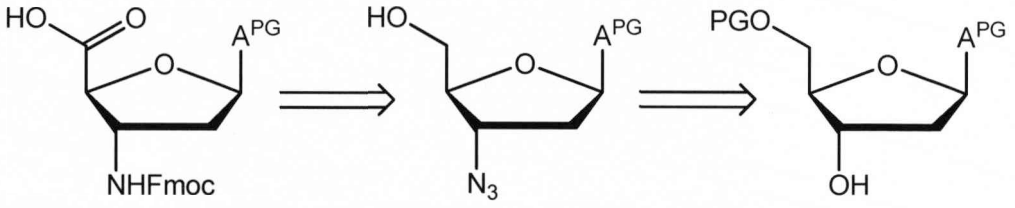


Figure 62: *2'-Deoxyadenosine derived peptide.*

2.4 2'-Deoxyadenosine-derived β-Amino Acid Synthesis

The synthesis of the analogous 2', 3'-dideoxyadenosine Fmoc amino acid was designed to follow a similar general strategy to that of the thymidine compound (see *fig. 63*). As with thymidine, the first step was protection of the 5'-OH. However, the exocyclic amine of adenine also required masking to prevent reaction at this site,

not only through production of the monomer, but also through the course of solid-phase peptide synthesis (SPPS).

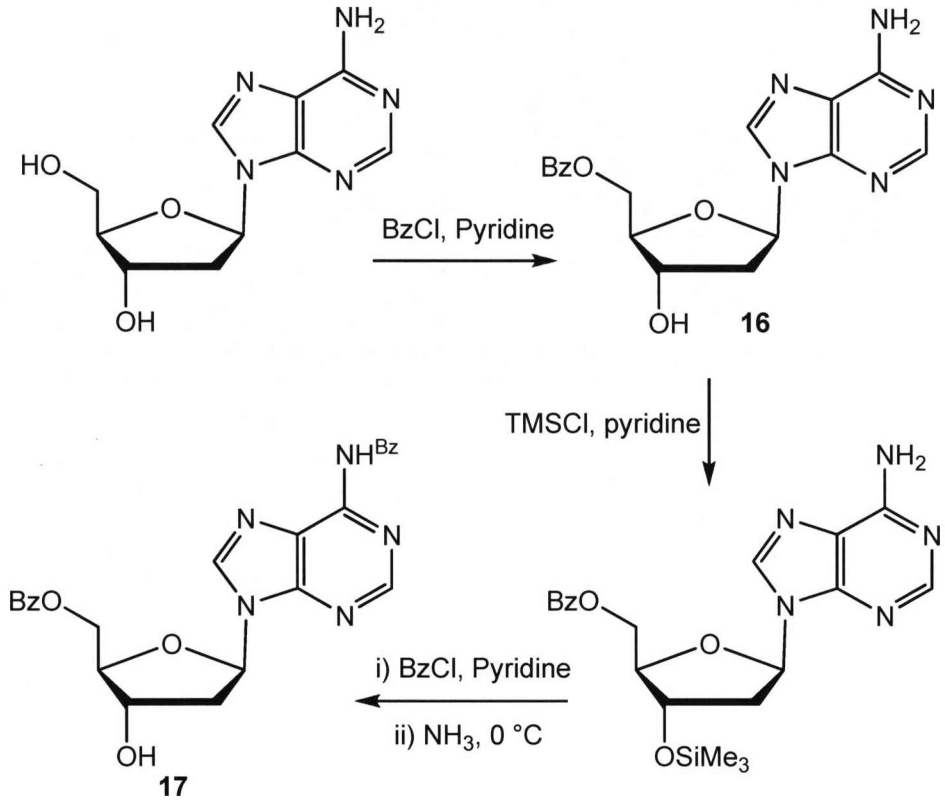


A = Adenin-9-yl

Figure 63

For this purpose, an orthogonal, acid-labile protecting group such as trityl or *bis*-*N*-Boc[227] may have been appropriate for the exocyclic amine to enable the concurrent cleavage of the peptide from the resin and removal of protecting group from the heterocyclic base. This is a common tactic employed in the synthesis of PNA.[209] Given then problems experienced with this type of procedure during attempts to make lysine appended thymidine oligomers, it was felt that a protecting group that could be removed whilst the peptide remained on the resin would carry a higher probability of successful characterisation of the peptide than an orthogonal method. Other combinations of protecting groups used for PNA synthesis include Boc with N-benzyl[118] and Fmoc with N-benzyl.[228] In each case Boc or Fmoc masks the peptidyl amine whilst benzyl protects the exocyclic amine of the purine or pyrimidine base.

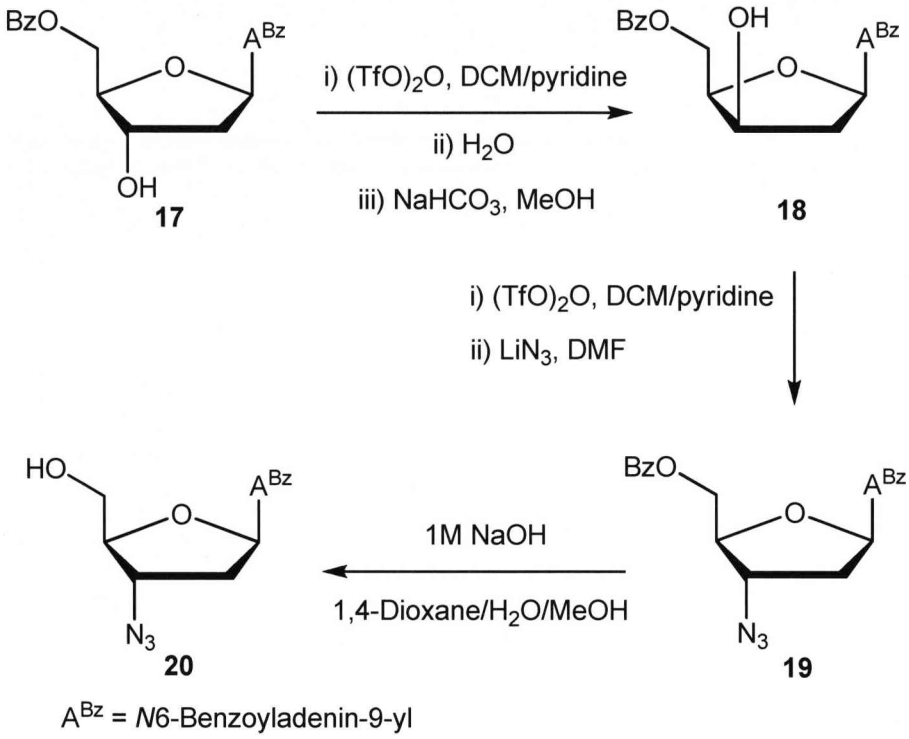
The use of Fmoc in conjunction with acyl groups is also an established protecting strategy in PNA synthesis[229] and appeared to be well suited to the synthesis of the target amino acid. It has been observed that N-acyl groups such as benzoyl offer stability under conditions commonly used for Fmoc SPPS.[230] It is also known that a benzoyl protecting group can be selectively removed from the sugar hydroxyl groups on a nucleoside in the presence of an *N*-6-benzoyl group, without significant cleavage of the *N*-6-benzoyl itself.[231] It was decided, therefore, that benzoylation of both the 5'-OH and exocyclic amine offered the simplest route to an appropriately protected monomer for both synthesis of the target Fmoc amino acid and SPPS. After protection of the reactive sites, two inversions of 3'-stereochemistry were required in a similar manner to that performed on thymidine to allow insertion of an azide with stereocontrol. As with the thymidine synthesis, after the introduction of azide, reduction, protection and oxidation were to follow to give the desired compound.



Scheme 5

The key dibenzoylated intermediate (**17**, *Scheme 5*) was easily accessed from 2'-deoxyadenosine *via* either a two-step process, wherein 5'-*O*-benzoyl-2'-deoxyadenosine (**16**) was isolated before further benzoylation of the amine,[232] or by a more convenient one-pot method which directly afforded the di-protected compound[233] (see *scheme 5*). In practical terms, the latter procedure was superior as the need to isolate and purify (**16**), which had a particularly difficult jelly-like physical form, was avoided. Thus 2'-deoxyadenosine in anhydrous pyridine was treated drop wise with a solution of 1.1 equivalents of BzCl also in anhydrous pyridine to introduce the 5'-*O*-benzoyl protecting group. As the exocyclic amine cannot be selectively benzoylated in the presence of the 3'-OH, the secondary hydroxyl group was then transiently protected with TMSCl before further treatment with BzCl afforded the fully protected nucleoside. The ability to selectively remove the silyl protecting group with base in the presence of the *N*-benzoyl moiety was now demonstrated. The silyl group was removed with ammonia solution at 0 °C to give the desired di-benzoylated product (**17**) in yields of greater than 80 %. NMR experiments run in d_6-DMSO showed an OH peak at 5.53 ppm which was confirmed by COSY to be the 3'-OH.

The change from a pyrimidine to a purine nucleoside also altered the approach to the S_N2 inversions for inserting the 3'-azide functionality. Neighbouring group participation from adenine could not be used in the same manner as thymine to create a 2,3'-anhydro species and alternative methods had to be explored.



A^{Bz} = *N*6-Benzoyladenin-9-yl

Scheme 6

Techniques in the literature are mainly reliant on a 3'-*O*-sulfonyl strategy similar to the one earlier disregarded when considering inversion of the thymidine 3'-OH group. Both mesylate[234] and trifluoromethanesulfonate[235] are popular derivatives for this type of S_N2 displacement reaction and both have remained so. Whilst mesylate is the poorer leaving group and requires longer reaction times, sources of triflate such as triflic anhydride are moisture sensitive and consequently more difficult to handle in practical terms. One notable exception to the use of sulfonate leaving groups is Gosselin and co-workers' modified Mitsunobu-type procedure. This is performed using diethylazodicarboxylate and PPh_3 with benzoic acid under conditions similar to those used to make 2,3'-anhydro thymidine derivatives.[236] After hydrolysis of the resulting benzyl ester, this can be repeated to introduce azide with diphenyl phosphorazidate or hydrazoic acid.[237]

Herdewijn's route to 3'-azido-2',3'-dideoxyadenosine using triflic anhydride[238] (see *scheme 6*) was chosen in preference to routes involving mesylate as it was believed that the benefit of the better leaving group outweighed the potential practical disadvantages. After first reacting the protected nucleoside with 1.5 equivalents of triflic anhydride in anhydrous DCM/pyridine solution, water was added and the intermediate 3'-*xylo* compound (**18**) was obtained. In addition to the desired product, another compound of similar R_f could be seen by TLC analysis. This compound was identified as 3'-benzoyl-3'-*O*-*xylo*-2'-deoxyadenosine (**22**, see *fig. 64*) and was a result of undesired neighbouring group participation during the displacement reaction leading to some migration of the 5'-benzoyl group to the newly inverted 3'-OH.[238] Although the two compounds were close on the TLC plate and had identical mass spectra and microanalysis compositions, distinguishing the two compounds by NMR was relatively easy.

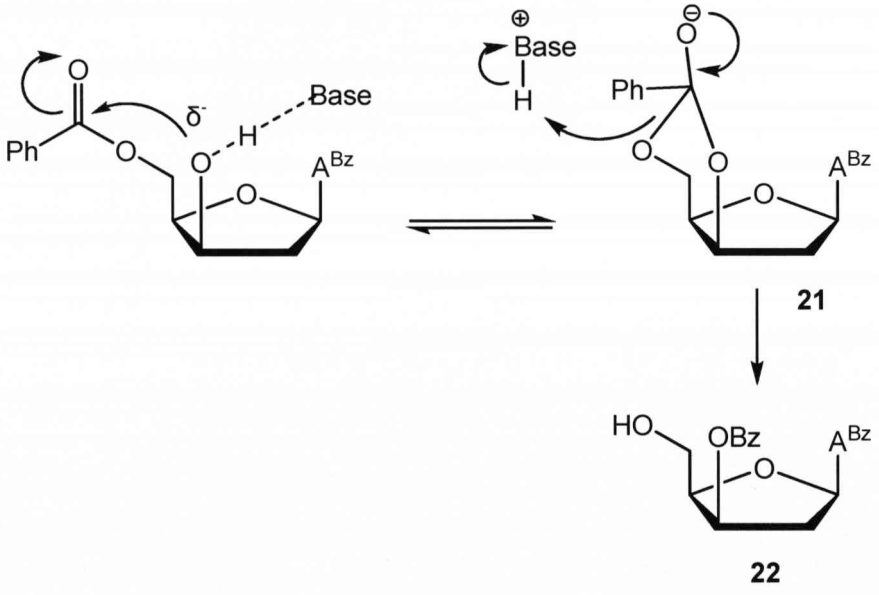


Figure 64: *Mechanism of 5'-O – 3'-O benzoyl migration.*

The NMR spectra for the 3'-*xylo*-OH compound had a very similar appearance to that of the starting dibenzoyl material, only slight shifts in the H3' signal, from 4.64 ppm to 4.56 ppm, and a corresponding shift in the C3' signal, from 70.9 ppm to 69.7 ppm, were observed, however, the 3'-*xylo*-OBz compound had a very different spectrum. As expected, large variations in the chemical shift of H3', H5', C3' and C5' were evident providing a strong indication that migration of the benzoyl group from the 5'- to the 3'-position had indeed occurred. Running the NMR experiments in *d6*-DMSO revealed strong OH signals which were prominent enough

to show COSY connectivity. The OH peak visible in both spectra could easily be assigned as either 5'- or 3'- due to COSY interactions, conclusively proving the identity of each of the compounds.

The 3'-*O*-benzoate is preferentially formed as the primary alcohol is a better leaving group than the secondary. Upon formation of the intermediate (**21**) it is more likely that the 5'-OH will depart resulting in isomerisation of the benzoyl group to the 3'- position. Two solutions to the problem of benzoyl group migration were possible. The first was to add a single equivalent of benzoic acid to the reaction mixture in order to preferentially protonate the 3'-hydroxyl group and prevent (**21**) from forming.

An alternative method was to allow the formation of the 3'-*O*-benzoate then isomerise back to the 5'-benzoate using a weak base such as sodium hydrogen carbonate either *in situ* or after the completion of the reaction (see *fig. 65*). The latter method repeatedly produced a superior yield of the desired isomer, approximately 50 %, compared to half that for the benzoic acid procedure.

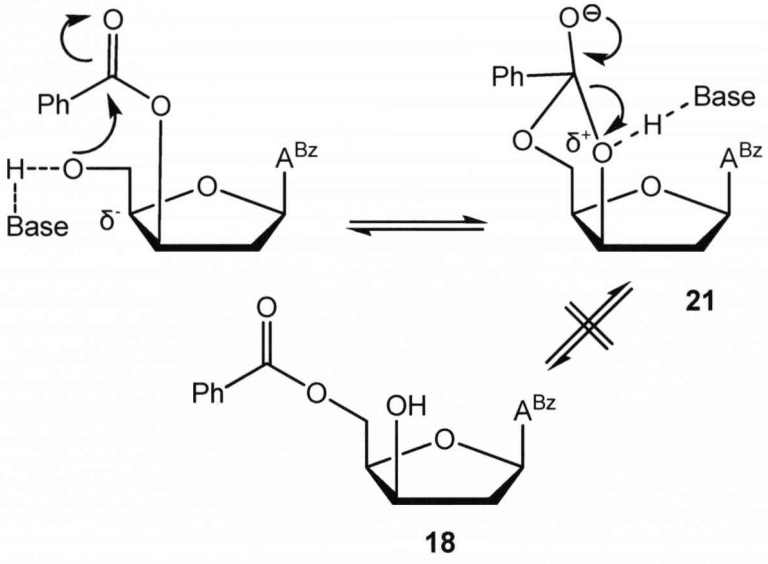
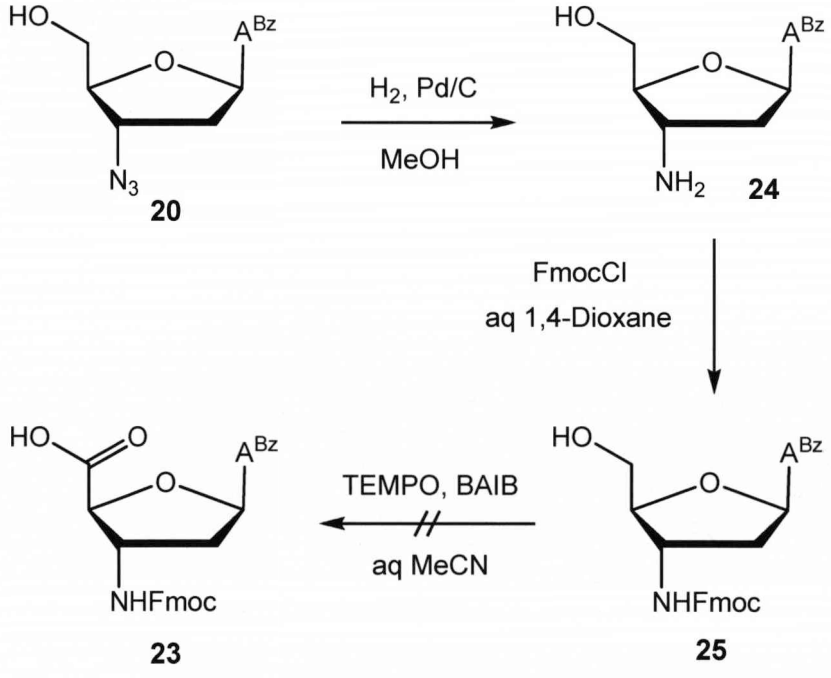


Figure 65: *Mechanism of 3'-O – 5'-O benzoyl isomerisation.*

The oily residue of the reaction of (**17**) with triflic anhydride was re-dissolved in methanol containing half an equivalent of $NaHCO_3$. The driving force behind the isomerisation reaction is the improved steric stability of the 5'-*O*-benzoate over the 3'-*O*-benzoate. The use of $NaHCO_3$ catalyses the reformation of (**21**) and the sterically less hindered 5'- isomer now prevails. The slightly less acidic secondary 3'-alcohol is more likely to be protonated in the presence of a weak base, preventing the establishment of an equilibrium back towards (**21**) and the 3'-benzoate.

The displacement of the new 3'-*xylo* hydroxyl group was achieved with a repetition of the previous inversion reaction conditions; 1.5 equivalents of triflic anhydride in anhydrous DCM/pyridine, this time using LiN_3 as the nucleophile instead of water. The fully protected azido nucleoside (**19**, *scheme 6, p. 76*) was obtained in an 86 % yield and, as with AZT, the compound was easily identified by the characteristic azide absorption in the IR spectrum at 2102 cm^{-1}. 5'-OBz deprotection was performed using NaOH in a mixture of water, 1,4-dioxane and methanol at 0 °C.[231] Strong base is selective for the removal of the *O*-benzoyl over the *N*-benzoyl as the base deprotonates at *N*-6-benzoyl leaving a delocalised negative charge at the amide. This negative charge then dramatically slows the rate of base-catalysed hydrolysis at this site allowing for the selective removal of the benzoyl group on oxygen, where this deprotonation process is not possible. Recrystallisation from hot methanol gave the desired azido alcohol (**20**) in 86 % yield.

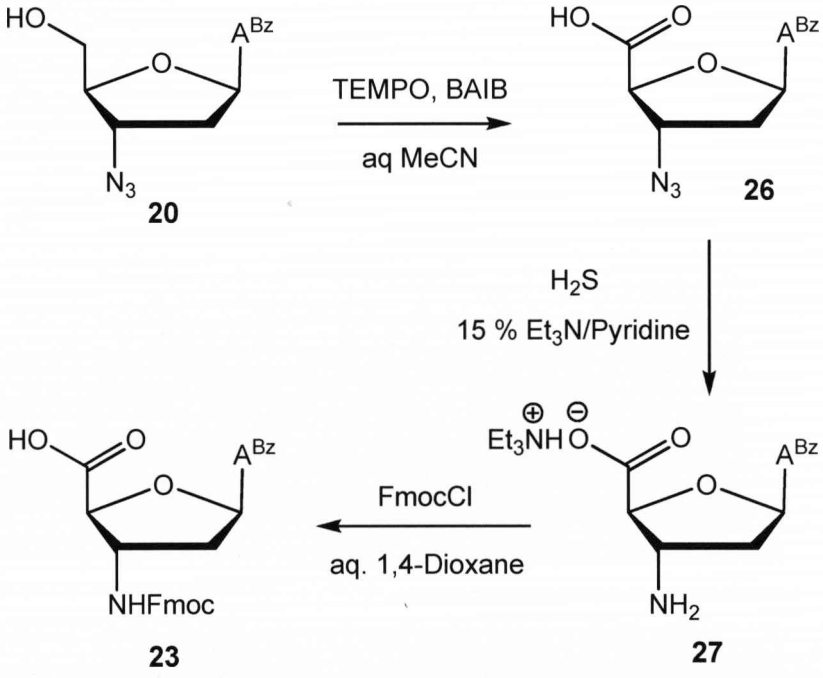


Scheme 7

At this stage it was proposed to follow the process that used to produce the thymidine Fmoc amino acid and reduce the azide by hydrogenation, followed by Fmoc protection and oxidation of 5'-C with TEMPO/BAIB to give the target amino acid (**23**, see *scheme 7*). The hydrogenation of the analogous AZT was carried out in methanol; however, (**20**) crystallised from methanol and could not be converted to

the amine in this solvent, as the product was totally insoluble. A simple solution was to add water to the methanol. Whilst this did not solvate the starting material, the product solubility was much improved and the hydrogenation proceeded smoothly affording the amino nucleoside (**24**) in quantitative yield. Successful reduction was confirmed by the shift of H3' 1 ppm upfield from 4.7 ppm in the azido nucleoside (**20**) to 3.7 ppm in amino alcohol (**24**).

Fmoc protection of (**24**) was successful as this reaction also required a water/organic solvent mix of 30 % aqueous 1,4-dioxane. Again, successful reaction was monitored by the shift of H3' to 4.4 ppm as well as appearance of Fmoc peaks at 4.25 and 4.38 ppm and the characteristic Fmoc aromatic pattern between 7.3 – 7.9 ppm. Unfortunately, the additional hydrophobicity introduced by the Fmoc group rendered the Fmoc alcohol (**25**) insoluble in aqueous acetonitrile, the solvent system required to oxidise the 5'-carbon with TEMPO/BAIB. Consequently, the oxidation with TEMPO/BAIB failed.



Scheme 8

Oxidation with TEMPO after reduction of the azido group was not possible as (**24**) was not sufficiently soluble in aqueous acetonitrile for the oxidation to occur. The poor solubility of the 3'-amino compound in aqueous solvents is likely to be due to the more aromatic character of adenine compared to thymine and this effect is exacerbated by the benzoyl protecting group of the exocyclic amine making the

heterocyclic base even more hydrophobic. The final option was to oxidise (**20**) before reduction,[239] producing the free amino acid (**26**, see *scheme 8*) which would subsequently be protected.

Oxidation proceeded smoothly using the standard conditions employed during thymidine oxidation[195] to give (**26**) in 82 % yield. The azido acid was easily characterised by the absence of the H5' protons in the ^{1}H NMR and shift of C5' to 171.4 ppm. The improved oxidation yield achieved with dA may lend support to the theory that the poorer crop of oxidised thymidine amino acid may be due to cleavage of the thymine double bond by BAIB[205]. With a much more aromatic character, adenine could be less prone to this type of oxidative cleavage of double bonds and hence the improved conversion rate. The improved yield alone is not sufficient to prove or disprove such a hypothesis.

Hydrogenation as a method of reduction for the azido nucleoside (**26**) proved unreliable, again due to solubility issues with both reactant and product. Combinations of water with methanol or ethanol were unsuitable. Addition of small quantities of dilute HCl,[240] in an attempt to induce salt formation and increase the water solubility of the product, were unsuccessful. There are several reports in the literature of direct preparation of carbamates from azides using transition-metal catalysis[241] or trimethylphosphine[242] and corresponding carbamoyl chlorides. These rely on attack of the metal or phosphine on the chloroformate and producing a species similar to a Grignard reagent or phosphorus ylide. Attack of this intermediate on the azide then liberates N_2 and subsequent hydrolysis gives the desired chloroformate. Attempts to utilise zinc powder[241] with the azido acid (**26**) did not result in any recognisable product. Some of the more unusual methods of reduction of azides to amines include radical reactions with AIBN/Bu_3Sn, $InCl_3$[243] or Fe with $AlCl_3$ or $BiCl_3$[244] but by far the most popular alternative to hydrogenation for nucleosides is the Staudinger reaction.[181] Triphenylphosphine attacks the azide forming an iminophosphorane then hydrolysis of this intermediate gives the desired amine and triphenylphosphine oxide.

The attractive aspect from our point of view was that there are many examples of Staudinger reductions with nucleosides in the literature performed in a variety of solvents.[245] For application to a compound with difficult solubility, this was a promising possibility. Conversely, the biggest drawback to the Staudinger reaction is the production of triphenylphosphine oxide. This is not only the natural driving force for the reaction but is also notoriously difficult to remove from reaction

mixtures. With already potentially difficult purification of the nucleoside in prospect, this was not desirable. However, resin bound triphenylphosphine[246] or scavengers[247] offered prospective solutions for purification. With this in mind Staudinger conditions were tested for reduction of azido acid (**26**). Although formation of the iminophosphorane was extremely rapid as judged by TLC, the intermediate was stable to hydrolysis with water and with aqueous ammonia to the extent where partial debenzoylation of the exocyclic amine occurred. This extended stability of the intermediate iminophosphorane prevented the Staudinger method from being a suitable reduction procedure and further alternatives for the azide conversion were sought.

1,3-Propanedithiol[248] and hydrogen sulfide[155,182,249] are both known to reduce azides to amines in a relatively clean and efficient manner. Practically, their popularity is limited due to strong odours, especially in the case of 1,3-propanedithiol. For this particular reaction H_2S appeared to be a good option for several reasons. Firstly, comparing the mechanism with that of 1,3-propanedithiol (see *figs. 66 & 67*), it can be seen how only elemental sulfur is produced as a by-product. Combined with performing the reaction in Et_3N/pyridine[155], the primary product would be the triethylamine salt of the amino acid, which would be water soluble and easily separated from the sulfur by-product by water washes.

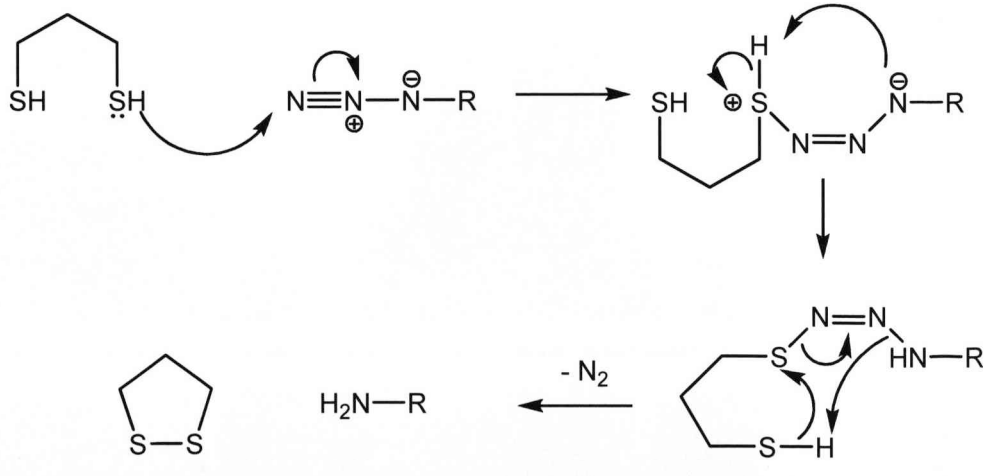


Figure 66: *Mechanism of propane1,3-dithiol mediated reduction of an azide.*

Lastly, directly obtaining a salt would prove useful in the final protection reaction with FmocCl as the salt would inevitably be required to solubilise the amino acid for this reaction to take place. The triethylamine salt (**27**) was indeed obtained after reaction with H_2S in 15 % Et_3N/pyridine. Along with large upfield chemical shift

changes at H3' and C3', 5.06 ppm to 4.04 ppm and 63.5 ppm to 55.0 respectively, triethylammonium peaks were also present in both spectra.

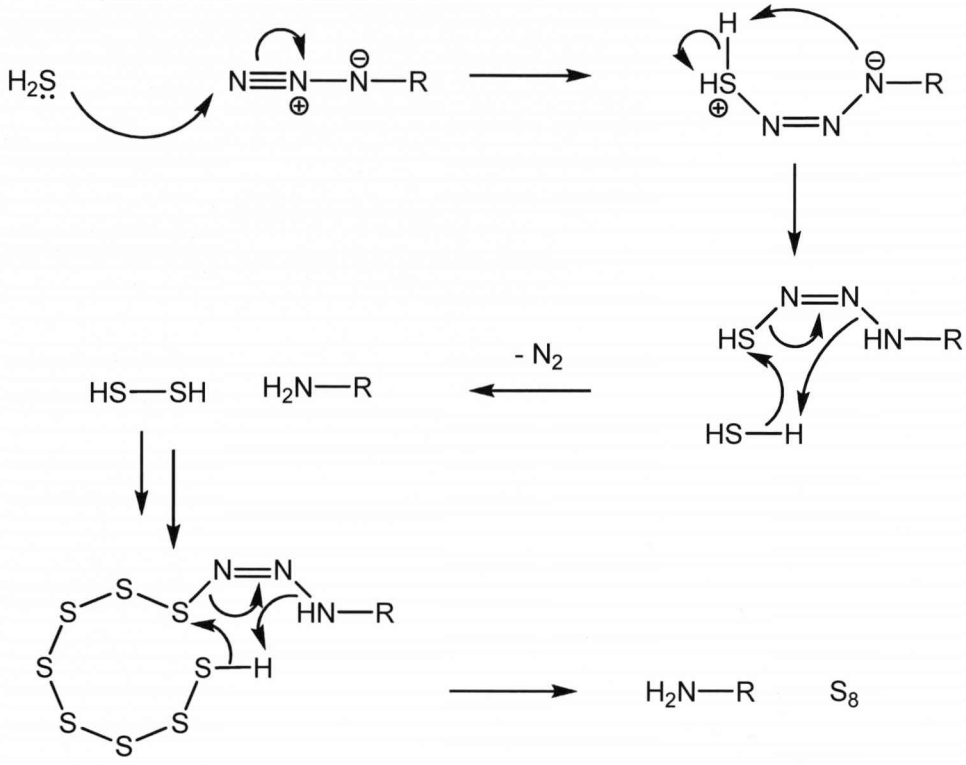


Figure 67: *Mechanism of H_2S mediated reduction of an azide.*

An interesting feature of the ^1H NMR spectrum of compound (**26**) was a significant downfield shift of the adenine H8 proton (see *fig. 68*). This implies a change in the sugar pucker from C3'-*exo* to C3'-*endo*, or "south" to "north". Electronegative substituents such as azide at C3' favour a south configuration where the electronegative group adopts an axial position.[250] Switching to a less electronegative group, as in this example converting an azide to an amine, releases this constraint and it can be seen that hydrogen bonding between H8 and the 5'-carbonyl would be possible, driving the sugar pucker to favour north (see *fig. 69*). This effect was not observed upon reduction of azido alcohol (**20**) to amino alcohol (**24**) and, therefore, must be linked to the presence of the 5'-carbonyl group. Favourable hydrogen bonding interactions are known to alter purine conformations. This behaviour is usually associated with guanine nucleotides,[6] although the H-8 proton of adenine has been observed to show similar hydrogen bonding interactions with the oxygen of the 5'-hydroxymethyl group.[20] This theory was further supported by the return of H8 to the chemical shift observed in the azido compound upon introduction of the Fmoc protecting group on the 3'-amine.

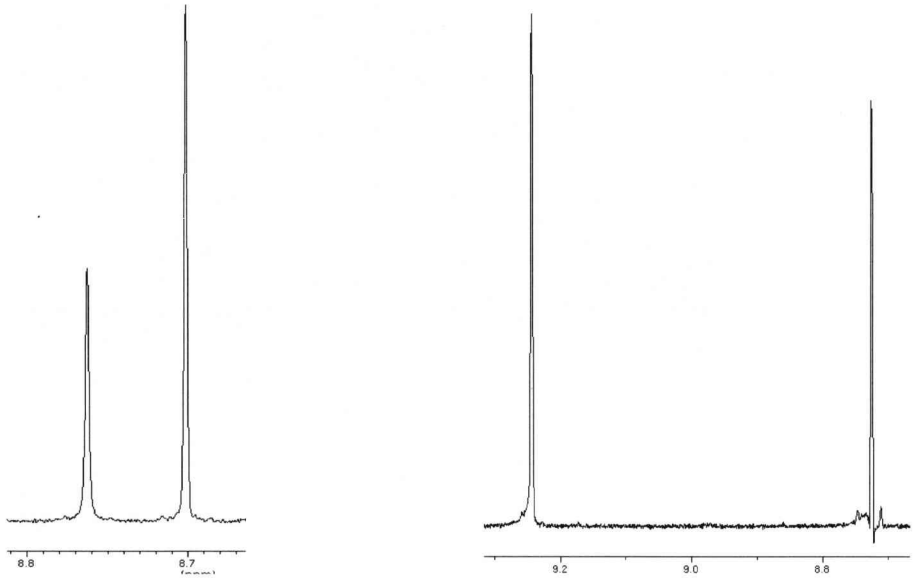


Figure 68: *NMR spectra of H8 for azide (**26**) (left) and amine (**27**) (right)*

Protection of the amine was achieved under the same conditions used for the thymidine analogue; a slight excess of FmocCl in aqueous 1,4-dioxane to afford (**23**) in 59 % yield. It was preferable to carry out the procedure at 0 °C to avoid di-protection, something not seen when protecting the thymidine compound. As for the protected thymidine amino acid, the successful addition of the Fmoc group was shown by the additional resonances present in the proton NMR between 7 and 8 ppm, assigned to the aromatic portion of the fluorenyl group, and at around 4.3 ppm, for the protons of the saturated 2-carbon chain.

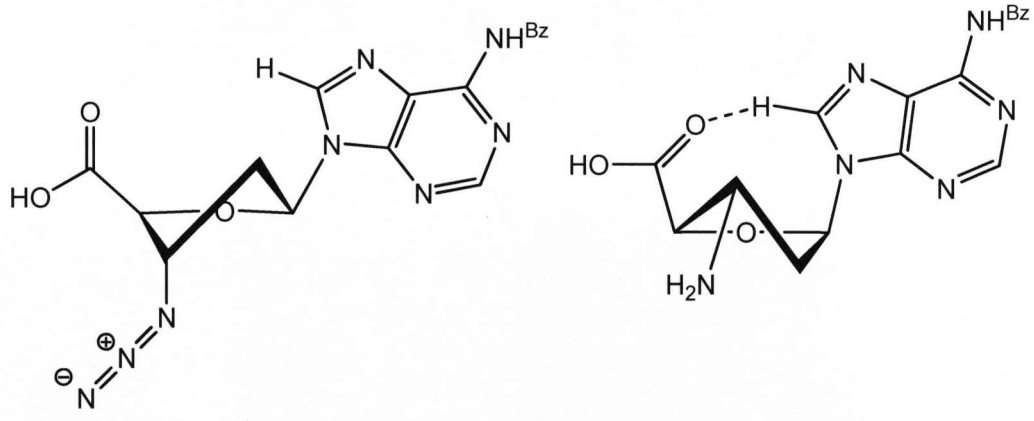


Figure 69: *Possible sugar pucker change upon reduction of 3'-azide group.*

With the protected amino acid in hand it was now possible to attempt to produce peptides similar to those made with the thymidine amino acid using the same SPPS techniques.

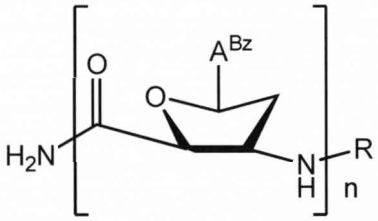
2.5 Solid-Phase Peptide Synthesis of 2'-Deoxyadenosine-derived β-Peptides

SPPS with the dA derived amino acid was initially performed using the optimised conditions developed for the analogous thymidine compound. Inter-residue couplings were reasonably efficient, at around 95 %. However, coupling of the 5'-terminal residue to the resin was much less effective, at 31 %. This is possibly due to the increased steric bulk of the protected dA amino acid compared to the thymidine compound (see *table 4*). Optimisation showed that the preferred solvent for the dA monomer was DMF and HBTU remained the preferred coupling reagent. Although this initial coupling was still relatively poor for SPPS, as the following couplings were highly efficient this was acceptable as a significant amount of peptide would be produced. Using this combination, successful peptide synthesis was achieved. Due to the extra hydrophobicity imparted by the *N*-6-benzoyl protecting groups, without removal of the 3'-terminal Fmoc moiety peptide (**28,** see *fig. 70*) proved too insoluble in aqueous solvents to be analysed by RP-HPLC. Characterisation of (**28**) by mass spectrometry was possible (see *table 5*).

Amino Acid[b]	Solvent	Coupling Agent	Coupling Yield/%
mdA1	DMAc	HBTU	31
mdA1	DMAc	PyBOP	49
mdA1	DMF	HBTU	74
mdA1	DMF	PyBOP	24
mdA2	DMF	HBTU	95

Table 4 [b] The entry "mdA1" refers to the first adenosine derived amino acid coupling to the resin and "mdA2" refers to the second adenosine derived amino acid coupling to mdA1.

After removal of the terminal Fmoc group, peptide (**29** see *fig. 70*) was successfully characterised by HPLC and ES-MS (see *table 5*). Attempts were then made to remove the *N*-6-benzoyl protecting groups. Applying the common conditions for benzoyl group deprotection of aqueous ammonia at 55 °C[251] resulted in complete degradation of the peptide.



(**28**) n = 4, R = Fmoc
(**29**) n = 4, R = H

Figure 70: *Nucleoside β-peptides derived from 2'-deoxyadenosine produced by SPPS.*

Oligomer	n	R	Molecular Formula	HR mass required	HR mass found
28	4	Fmoc	$C_{83}H_{71}N_{25}O_{14}^{2+}$	820.7806	820.7694
29	4	H	$C_{68}H_{61}N_{25}O_{12}^{2+}$ $C_{68}H_{62}N_{25}O_{12}^{3+}$	709.7460 473.4998	709.7610 473.5045

Table 5

This is a problem which is known in PNA synthesis, especially with proline derivatives, and is usually a consequence of intramolecular attack of the *N*-terminal amine on the carbonyl group of its nearest neighbour, forming a dihydropiperazine derivative (see *fig. 71*). This is then subject to further hydrolysis, eventually affording dipeptide units.[252] Given the relative stereochemistry of the nucleoside β-amino acids, this pathway is an improbable route for decomposition as the eight-membered piperazine intermediate would be extremely strained and unlikely to form. This was confirmed by the absence of any dipeptide fragments in the reaction liquors.

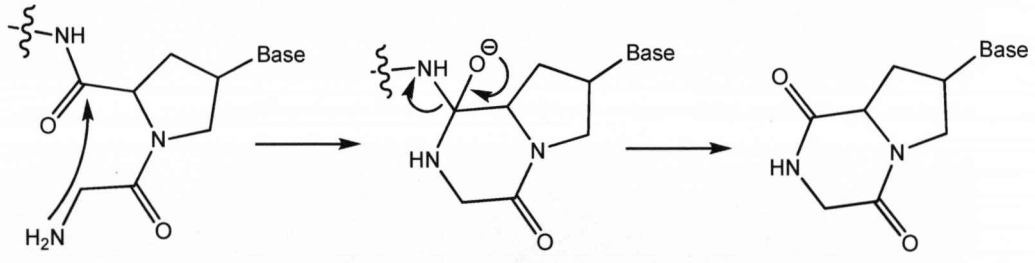
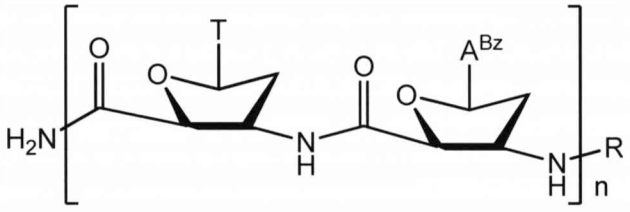


Figure 71:

An alternative, milder procedure to heating the peptide in ammonia solution is to dilute the ammonia solution with 50 % 1,4-dioxane as a co-solvent and perform the reaction at ambient temperature.[253] Under these conditions some deprotection was observed by mass spectrometry. However, more than 10 other major peaks

were detected by HPLC leading to the conclusion ammonia was not a suitable reagent for the removal of the benzoyl groups. NaOH in aqueous methanol[254] also produced a similar HPLC trace but no deprotection products were detected by mass spectrometry.

Using the optimum conditions developed for the synthesis of both the T and dA derived peptides a mixed sequence of the form TATA was successfully produced in Fmoc on (**30**) and off (**31**) forms (see *fig. 71*) and characterised by ES-MS (see *table 6*).



(**30**) n = 4, R = Fmoc
(**31**) n = 4, R = H

Figure 71: *TATA mixed sequence peptides.*

Oligomer	Sequence	R	Molecular Formula	HR mass required	HR mass found
30	TATA	Fmoc	$C_{69}H_{65}N_{19}O_{16}^{2+}$	707.7423	707.7449
31	TATA	H	$C_{54}H_{55}N_{19}O_{14}^{2+}$	596.7083	596.7112

Table 6

Although different solvents were preferred by each nucleoside, T preferring DMAc whilst dA coupled best in DMF, switching between solvents during the synthesis was possible without greatly affecting coupling yields. This indicated that solvent interactions of the uncoupled monomer have a larger influence on coupling efficiency than solvation of the resin bound peptide. In the absence of viable deprotection conditions for the adenine *N*-6 benzoyl group, structural characterisation of the dA tetramer was attempted with the protected oligomer.

2.6 Structural Characterisation of 2'-Deoxyadenosine-derived β-Peptides

As it was not possible to remove the benzoyl protecting groups without denaturing the peptide chain, structural characterisation of protected peptides (**28**) and (**29**) was attempted with the same set of NMR experiments used for the thymidine derived oligomers; DFQ-COSY, NOSEY, TOCSY and VT-NMR. Unfortunately, analysis of the NMR spectra obtained showed that, unlike the thymidine derived oligomers, the inter-residue amide protons were poorly resolved

and also coincided in chemical shift with those of the benzoyl protecting groups. This made measurement of temperature coefficients extremely problematic and observation of nOe couplings was not possible. Therefore, the existence of structure within the protected dA derived peptides could not be determined by NMR.

2.7 Conclusion

Fmoc β-amino acids derived from both thymidine and 2'-deoxyadenosine were successfully synthesised from the parent nucleosides. The T amino acid was produced in 6 steps with an overall yield of 18 % from thymidine and the dA amino acid obtained in 6 % yield over 7 steps from 2'-deoxythymidine. Both amino acids were polymerised using optimised Fmoc solid-phase peptide synthesis protocols affording homo-oligomers of varying lengths and a tetrameric mixed sequence of the form TATA. A tetrameric thymidine derived homo-oligomer was examined in d_6-DMSO solution by NMR (as the peptides were insufficiently soluble in $CDCl_3$ or MeOD). This peptide was found to adopt an 8 helical conformation with the sugars occupying a zigzag arrangement at the centre of the structure and the bases presented in two planes on opposite sides of the helix. UV melting studies appeared to corroborate the findings of the NMR experiments and showed that single strands of β-peptide produced an almost sigmoidal melting curve consistent with denaturation of ordered structure. It is still uncertain as to whether a thymidine-derived homo-oligomer successfully annealed with a complementary homo-dA DNA strand as this melting experiment was inconclusive.

Recently, Chandrasekar and co-workers have undertaken independent studies on the same thymidine derived oligomers produced using Boc-based solution-phase chemistry.[160] Interestingly no significant issues of cleavage of the glycosylic bond under acidic deprotection conditions were reported. Structural analysis was performed in a mixture of MeOD/$CDCl_3$ and resulted in assignment of the solution structure as a 12-helix, analogous to that of Gellman's *trans*-(ACPC).[160] One major consistency with the current work was noted. As could be expected, both models of the tetramer placed the sterically demanding thymine towards the outside of the helix. The difference in NMR solvent is important as, even though DMSO is often the solvent chosen for NMR experiments with short peptides, changing solvents can produce quite different spectra for the same β-peptide compound.[89] The possibility that a highly functionalised oligomer such as a nucleoside β-peptide adopts solution conformations which are greatly solvent dependent cannot be ruled out without further investigation.

It was not possible to assign any structure to the homo-oligomers derived from 2'-deoxyadenosine due to the presence of protecting groups. All attempts to remove the benzoyl moieties resulted in degradation of the peptide chain and, therefore, structural characterisation was attempted on the protected oligomers. Overlap of the benzoyl aromatic resonances with those from the inter-residue amide protons prevented the observation of through space connections and measurement of temperature coefficients which were key to assigning the structure of the thymidine derived oligomers. It is hoped that future work with an alternative *N6*-protecting group will facilitate structural analysis.

Chapter 3

Results and Discussion 2:

Fluorous-Phase Synthesis

3 Results and Discussion 2

3.1 Introduction to Fluorous-Phase Synthesis

The technique of solid-phase synthesis for both peptides and oligonucleotides is limited by several factors; scalability, reduced reactivity at resin surfaces due to steric hindrance and lack of a reliable quantitative method of analysis of the coupling reactions. During SPPS for this project all of these drawbacks were evident. It was necessary to load the solid-phase resin at 25 % of its full capacity in order to prevent interchain aggregation as the polymers were extended. This had obvious implications for larger scale syntheses. As the nucleoside β-amino acid was also sterically bulky, an extended reaction time of 5 hours was required for coupling of the first residue to the resin and 30 minutes for each subsequent inter-residue coupling, significantly extending the synthetic cycle. The absence of compatible analytical methods for solid-phase was mainly an issue during optimisation of the Fmoc SPPS protocols. Colormetric tests such as the TNBS test were only suitable for giving a qualitative indication on the success or failure of any given coupling reaction. For any quantitative data it was necessary to either cleave the compound from the resin and analyse by UV or HPLC or to deprotect the Fmoc group and assay for dibenzofulvene. Both of these methods involved terminating the synthetic cycle and beginning the synthesis afresh. The combination of these factors made optimisation of the SPPS procedure extremely time consuming and labour intensive.

Whereas solid-phase procedures are subject to these types of constraint, solution-phase chemistry is commonly performed on industrial scale with full analysis possible at any point. It is clear that a reliable method of solution-phase synthesis for polypeptides would be attractive. Fluorous-phase synthesis (FPS) presents an interesting, solution-phase alternative to the traditional solid-phase approaches to oligomer production. It is known that short oligonucleotides, oligosaccharides and peptides can be synthesised in solution-phase *via* fluorous chemistry[255] and that the fluorous components used are generally recyclable.[256] Additionally, the chemically inert, non-flammable, thermally stable[257] and non-toxic[258] nature of the highly fluorinated solvents used makes this type of chemistry equally appealing with respect to environmental and safety considerations.

FPS utilises the unique solubility properties of highly fluorinated compounds as a handle for purification of organic compounds. The principle of FPS relies on the

fact that organic compounds containing a high proportion of fluorine are preferentially soluble in perfluorinated solvents (perfluoroalkanes, perfluorodialkyl ethers and perfluorotrialkyl amines) and these perfluorinated solvents have a very low miscibility with organic solvents.[259] Appending an organic compound with a removable perfluorinated tag provides a convenient method of purification of a product or removal of an unwanted by-product by simple partition between organic and fluorous phases.

This principle was first demonstrated in practice when Horváth and Rábai attempted to tackle the problems posed by separation of long chain aldehydes (C8 or more) from the rhodium/triphenylphosphine catalyst used in the oxo process.[260] One solution was to use an aqueous biphasic system where organic products were easily partitioned from the preferentially water-soluble catalyst. This approach was limited by the poor solubility of the starting alkenes in water. As the alkenes were more soluble in fluorous solvents it was possible to perform the reaction in a biphasic mix of toluene and C_7F_{14} in the presence of a specially designed, fluorous soluble Rh catalyst $HRh(CO)(P[CH_2CH_2(CF_2)_5CF_3]_3)_3$.[261] The product aldehydes were preferentially soluble in the organic phase and were easily separated from the catalyst, as this remained in the fluorous layer (see *fig.* 72). Since this initial demonstration of FPS, fluorous techniques focussed on liquid-liquid extraction have become known as "heavy" fluorous chemistry.[262]

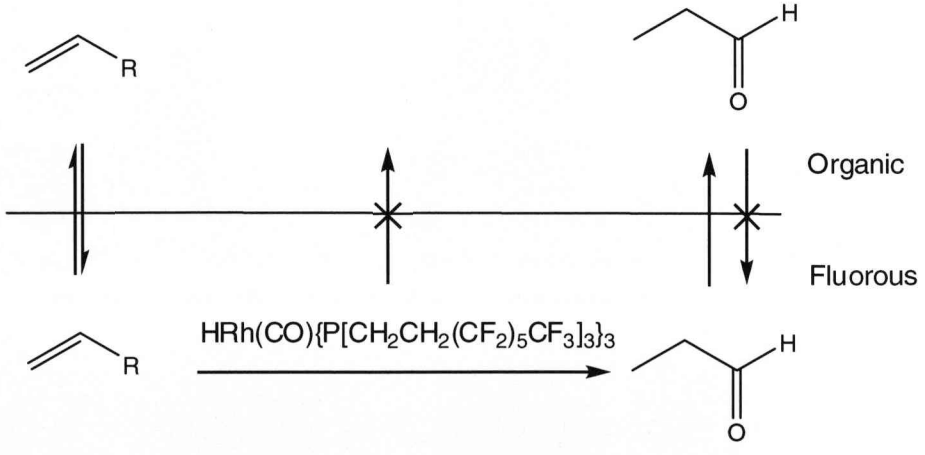


Figure 72: *Fluorous-phase oxo reaction.*

3.2 "Heavy" Fluorous Chemistry

Successful fluorous liquid-liquid extraction is dependent on the number and distribution of fluorines within a molecule. A fluorinated compound containing

greater than 60 % fluorine gives the best performance in terms of partition.[263] In addition, to protect the rest of the molecule from the powerful electron withdrawing effects of multiple fluorines, a hydrocarbon spacer must be present between the site of reaction and the perfluorinated chain. Usually $(CH_2)_2$ or longer, this organic soluble section must, as far as possible, be appropriately shielded from the fluorous solvent to maximise partition.[264] FC-72 (perfluorinated hexane) is the solvent of choice for extractive work-up with organic solvents as it is the least miscible of all the commonly used fluorous solvents,[265] ensuring the highest partition. Heavy fluorous methods (see *fig. 73*) have found extensive use in procedures such as the oxo process in which separation of one or more components of the reaction is problematic. A further example of this is the Mitsunobu reaction.

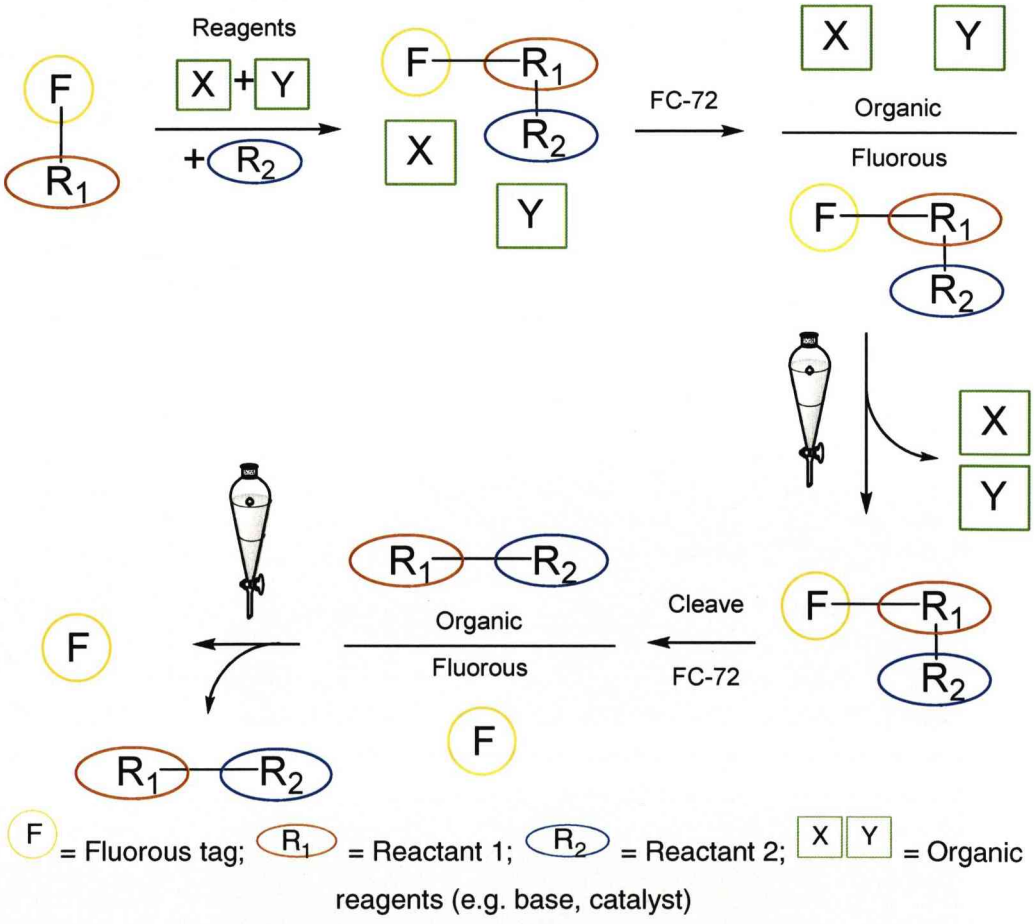


Figure 73: *Schematic of heavy fluorous-phase process.*

Mitsunobu reactions are synthetically very useful but the presence of phosphine oxides and hydrazine derivatives as by-products often complicates purification. Mitsunobu reactions have been carried out with fluorous modifications

to phosphine, azodicarboxylate and even the nucleophile, with excellent separation of the fluorous component.[266] The gallic acid derivative 3,4,5-tris(4-(perfluorooctyl)butoxy)benzoic acid (see *fig. 74*) displays all the requirements for good liquid-liquid extraction; 60.9 % fluorine with good shielding of the hydrocarbon segment and is also readily soluble in organic solvents such as THF.

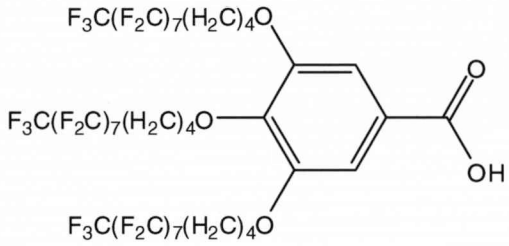


Figure 74: *Johansson's fluorous gallic acid derivative.*

This fluorous gallic acid derivative was successfully reacted in a Mitsunobu procedure with a number of primary and secondary alcohols in the presence of DIAD and PPh_3.[267] Isolation of the pure products using conventional liquid-liquid extraction was found to be inefficient for some of the perfluorinated esters due to the altered ratio of fluorous to organic content. In these cases, purification was achieved using an alternative technique known as organic liquid/fluorous solid separation. Similar to a standard organic crystallisation, a fluorous compound soluble in one organic solvent, in this instance THF, is precipitated out of solution by the addition of another solvent, such as 50 % MeOH/$CHCl_3$. Isolated yields for most of the alcohols tested exceeded 90 %.

3.3 "Light" Fluorous Chemistry

A distinct division emerged in the field of fluorous chemistry with the introduction of fluorous solid-phase extraction (F-SPE).[268] Whereas much of the early work such as that on the fluorous oxo reaction was "heavy", the new technique employed fluorinated silica to perform solid-phase extractions. Termed "light" fluorous chemistry, this method improved the scope of fluorous chemistry as the shorter fluorous tags used were much more soluble in common organic solvents. However, with less fluorination, these tags were now unsuitable for use in liquid-liquid extraction and separation *via* fluorous solid-phase was established as the purification method of choice.

F-SPE is, in essence, simplified column chromatography and as such has a standard and reverse phase (see *fig 75*).[269]

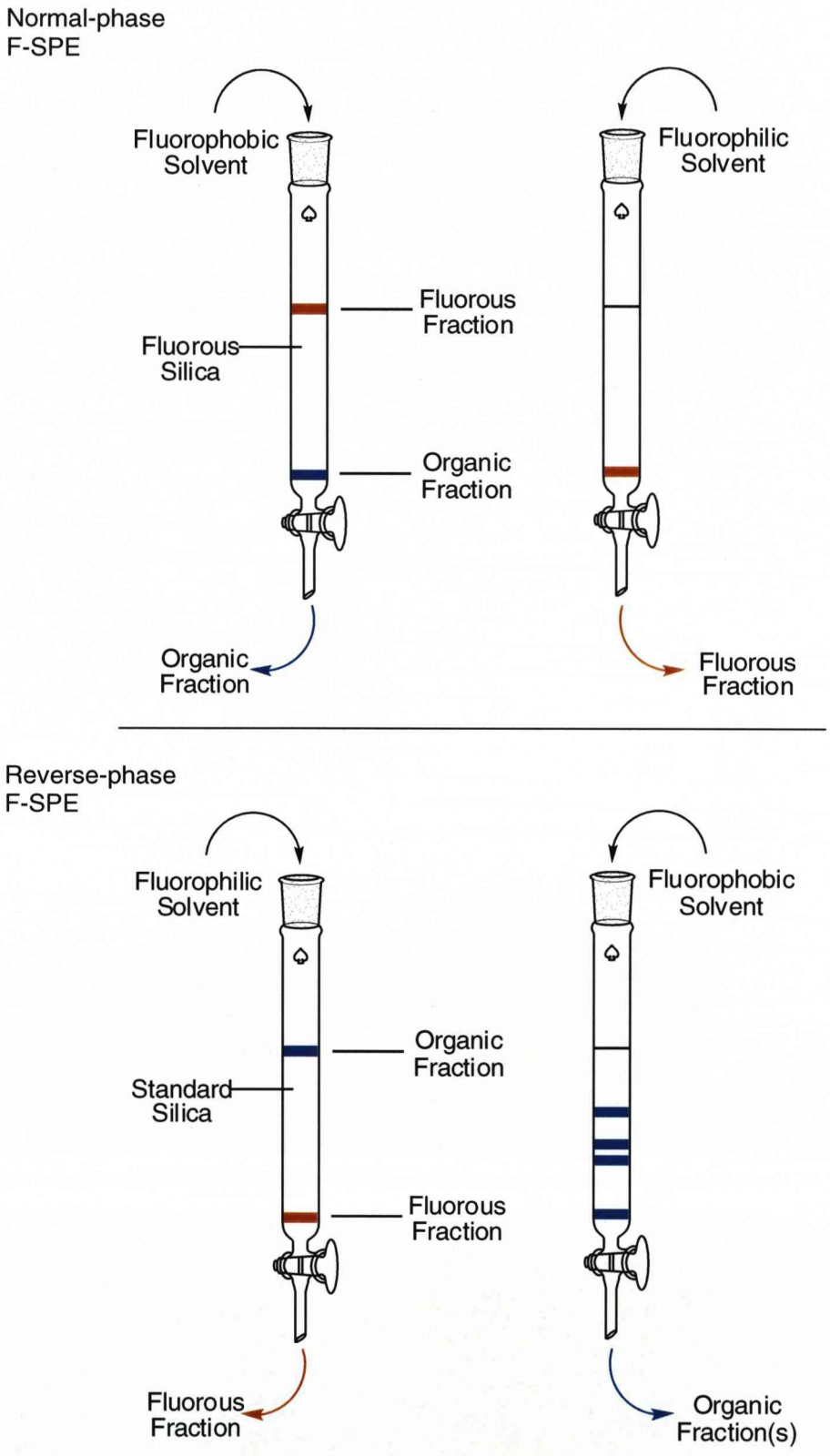


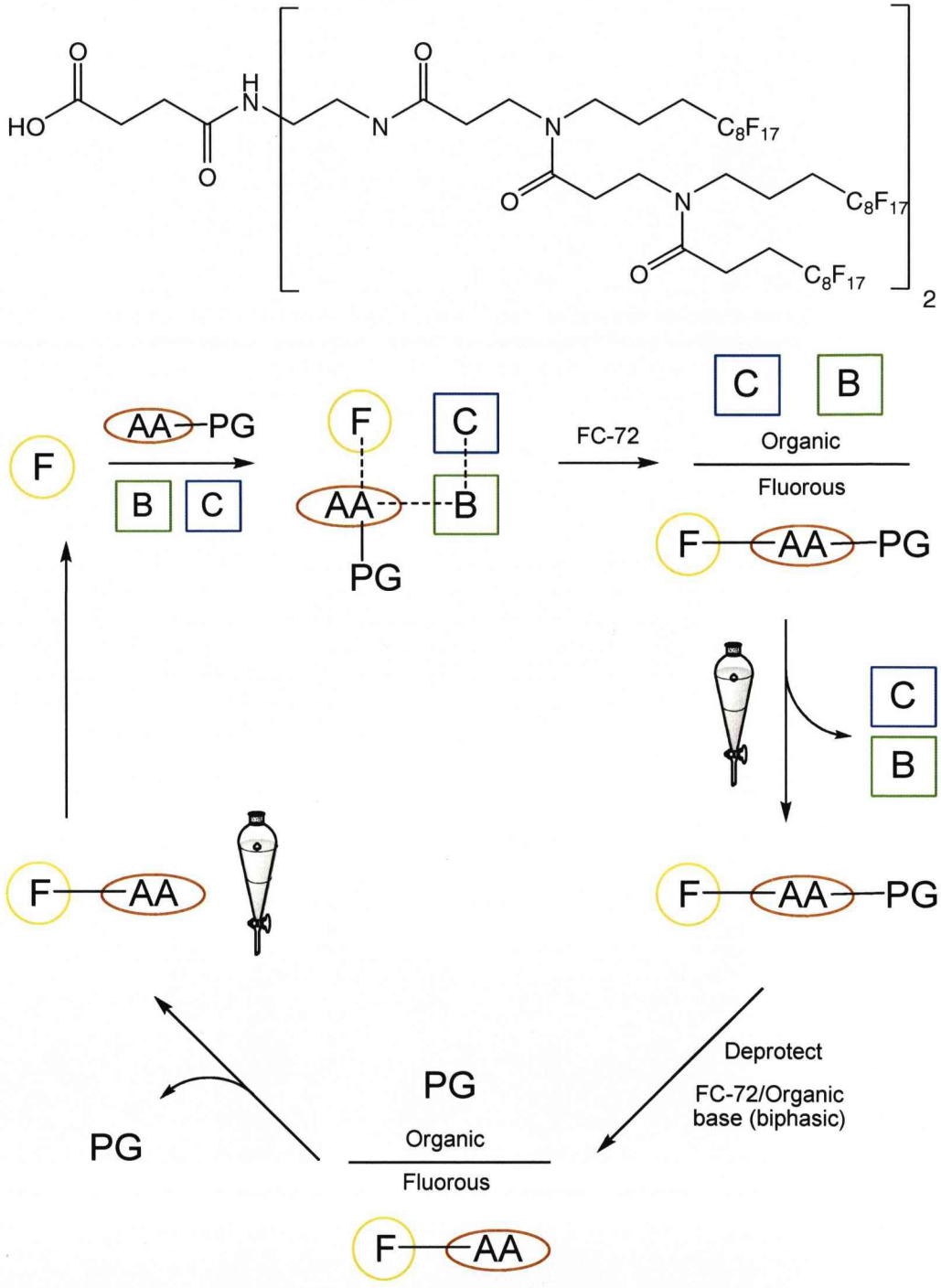
Figure 75: *Normal-phase F-SPE (top) and reverse-phase F-SPE (bottom). (Figure adapted from reference 262).*

In the standard method, the crude mixture containing both fluorous and organic fractions is loaded onto a silica stationary phase functionalised with perfluoroalkyl chains, known commercially as Fluoro*Flash*™. The column is then eluted first with a fluorophobic eluent, for example 80 % MeOH/H_2O, followed by a fluorophillic wash with an anhydrous organic solvent such as anhydrous methanol or acetonitrile. Water is considered the perfect fluorophobic solvent[270] and as a result the aqueous solvent will not elute any fluorinated compound from the silica. All organic compounds will, however, be eluted in this fraction leaving the fluorous compounds to be removed by the later fluorophillic eluent. In reverse-phase F-SPE, non-fluorinated silica is used and eluted first with a blend of fluorous and organic solvents. Reverse phase is mainly used for removal of fluorous catalysts or reagents from organic targets as, after the removal of the first fluorous fraction, conventional column chromatography can then be performed on the retained organic compounds if required.

3.4 Fluorous-Phase Peptide Synthesis (FPPS)

For each stage in the synthetic sequence of peptides, fluorous chemistry has been applied in some way. Fluorous capping reagents have been used to both remove failure sequences and isolate finished peptides from crude mixtures cleaved from solid-supports. Filippov and co-workers functionalised a 35 residue α-peptide constructed by SPPS with a fluorous methanesulfonyl ethoxycarbonyl group and purified the tagged chain by fluorous HPLC.[271] Montanari and Kumar capped deletion sequences with a simple perfluoroalkyl chain and separated these from the unmodified peptide by F-SPE.[272] A wealth of fluorous protecting groups and coupling reagents have recently become commercially available[273] reflecting the increasing interest in this type of separation technology. By far the most radical approach to peptide synthesis utilising fluorous reagents is the replacement of solid-phase resins with "fluorous supports".

Mizuno and co-workers synthesised the fluorous equivalent of Wang resin and also a trialkoxybenzhydryl and a *t*-butyl support from highly fluorinated compound hexakis-(fluorous chain) butanoic acid (Hfb-OH, see *fig. 76*) by reaction with 1,2-ethylenediamine and the appropriate linker.[274] The tripeptide thyrotropin releasing hormone was produced by standard Fmoc peptide synthesis protocols using a heavy fluorous approach on the trialkoxybenzhydryl support.



F = fluorous support; AA = Amino acid; C = Coupling agent; B = base; PG = protecting group

Figure 76: *Mizuno's Hfb-OH fluorous support and schematic of fluorous-phase peptide synthesis.*

The removal of the Fmoc group from the linker amine was achieved in a biphasic mix of FC-72 and 5 % piperidine in DMF. Dibenzofulvene was removed in the organic layer by simple separation of the two phases after the deprotection was complete. Coupling of amino acids was performed with PyBOP and DIPEA in a mixture of HFE-7200 (C_4F_9OEt) and CH_2Cl_2 and the finished peptide was cleaved from the fluorinated support with 2.5 % aqueous TFA. All intermediates were purified using an extractive work-up with FC-72 and the final product separated from the fluorous compound by extraction with water. Pentapeptide salusin-α was also synthesised using a similar trialkoxybenzhydryl support.[256] Even though both of these peptides still required purification by HPLC, the fact that production was possible and uncomplicated demonstrated the potential of fluorous chemistry to offer a real alternative to SPPS.

3.5 Towards Fluorous Synthesis of Nucleoside β-Peptides

Preliminary investigations into the application of heavy fluorous chemistry for synthesis for nucleoside β-peptides were conducted during a 12 week visit to the University of Tokyo. The fluorous support used in these experiments was the dendrimer FD-O(H) (**32**, see *fig. 77*) which is structurally reminiscent of Johansson's highly fluorinated gallic acid derivative.

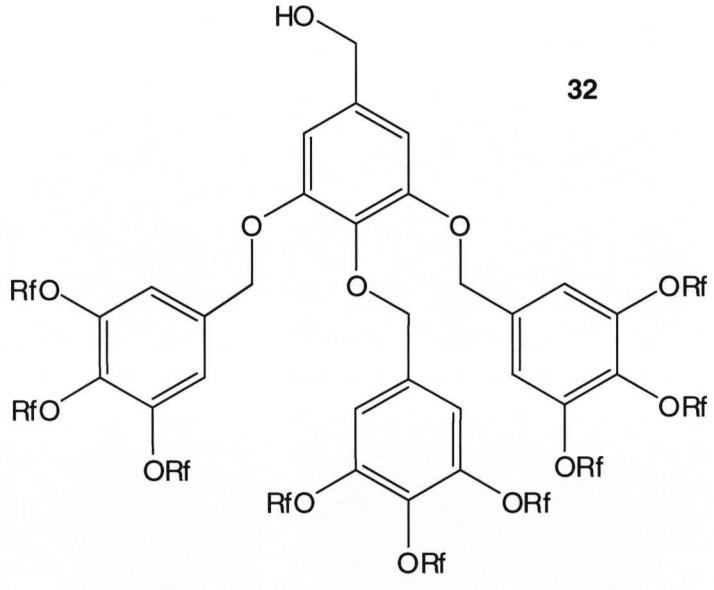
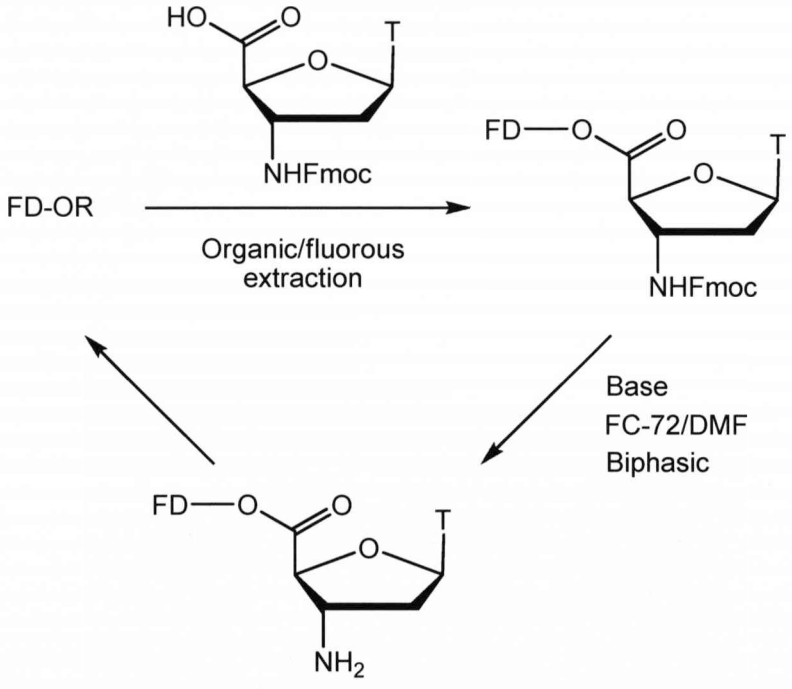


Figure 77: *Fluorous dendrimer (FD-O(H)).*

The fluorine content, 61.69 %, is above the ideal 60 % required for liquid-liquid extraction and the phenyl rings and three methylene groups separating the hydroxy group from the perfluorinated alkyl chains provide excellent insulation of the reaction site from the electron withdrawing fluorines. Roughly following the method of Mizuno *et al.*[274], it was proposed to couple the nucleoside β-amino acid with the fluorous dendrimer in a mixture of fluorous and organic solvents in the presence of an ester bond-forming reagent and base. After coupling, separation would be achieved by addition of FC-72 and extraction of excess reagents with water and other immiscible organic solvents. Deprotection could then be performed in a similar FC-72/DMF biphasic system in preparation for chain extension with the subsequent amino acid (see s*cheme 9*).

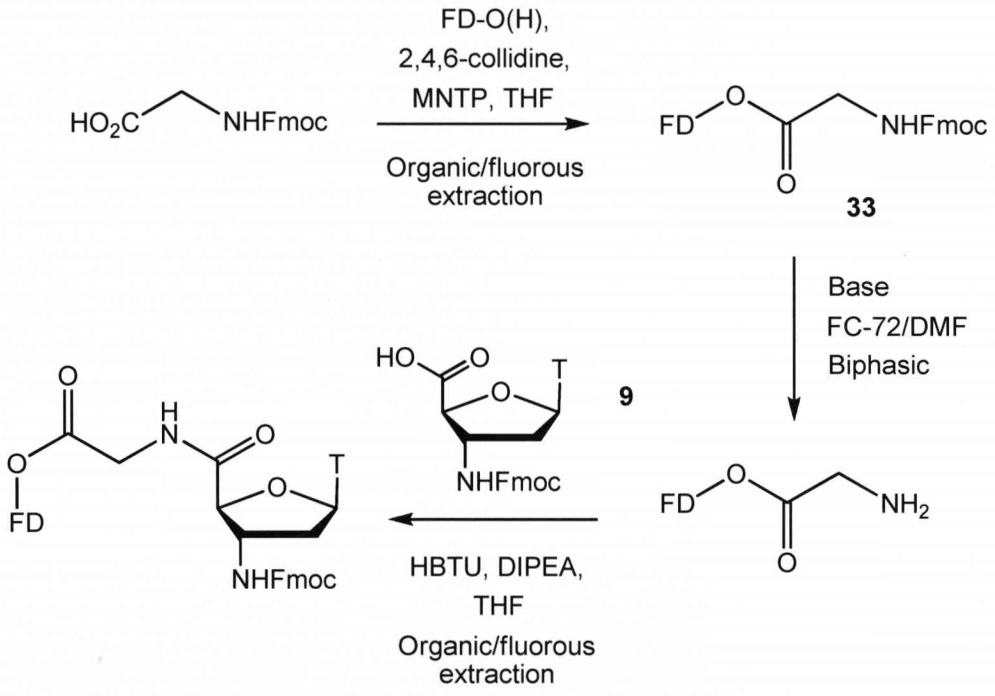


Scheme 9

After screening many permutations of temperature, coupling reagents and solvents, no suitable reaction conditions for the coupling of the amino acid could be identified. Earlier SPPS work had demonstrated that the nucleoside amino acid was approximately 30 % less efficient in forming ester bonds than amide linkages with solid-phase resins (see *section 2.2*). The failure of this esterification was thought to be due to severe steric hindrance from both nucleoside and dendrimer similar to that found at the surface of a resin. Drawing the analogy with SPPS it was decided

to attempt to attach a linker to the dendrimer in order to alleviate the problem of steric interference and facilitate coupling.

The ideal linker would not only be sterically small, in order to avoid the difficulties encountered by the large nucleoside, but would also offer a terminal amine group for condensation with the nucleoside amino acid instead of the alcohol moiety of the dendrimer. A smaller α-amino acid such as glycine would fit both criteria and with this in mind, attempts were made to couple Fmoc-Glycine-OH with the fluorous support (see *scheme 10*).



Scheme 10

Conducting the esterification reaction of Fmoc-Gly-OH with the dendrimer in solution allowed analysis of the mixture by TLC, something which is impossible with standard SPPS. Condensation was successfully accomplished using the ester bond forming reagent 1,3-dimethyl-2-(3-nitro-1,2,4-triazol-1-yl)-2-pyrrolidin-1-yl-1,3,2-diazaphospholidinium hexafluorophosphate (MNTP, see *fig. 78*)[275]. MNTP is a phosphonium coupling agent similar to PyBOP but incorporating a 3-nitro-1,2,4-triazole moiety instead of a hydroxybenztriazole and a bridged dimethyl-diaza functionality in place of two of the usual three pyrrolidine groups. The ester bond formed in under 1 hour in the presence of 5 equivalents of 2,4,6-collidine in THF at 45 °C. Work-up of the reaction mixture was simply a matter of addition of a small volume of FC-72 and washing the resulting fluorous layer with water and $CHCl_3$. As

a precaution against loss in the washings, both water and organic layers were then backwashed with FC-72. It is uncertain if this was absolutely necessary as the contents of the respective layers were not analysed.

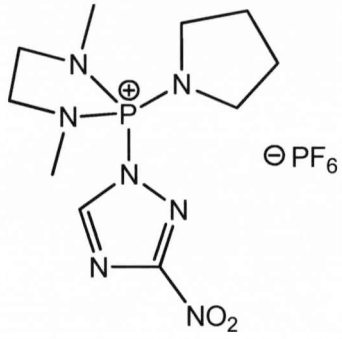


Figure 78: *MNTP*

Another major advantage of fluorous-phase synthesis in terms of analysis was then demonstrated. A sample of the residue after evaporation of the FC-72 was dissolved in a mixture of 3:1 $CDCl_3/C_6F_6$ and analysed by NMR spectroscopy. After this data was obtained, the mixed NMR solvents could then be removed under reduced pressure, the sample returned to the bulk material and the synthetic procedure continued without loss of any material. The NMR spectrum showed a nearly identical resonance pattern in the aromatic region to that observed for the Fmoc group of Fmoc nucleoside β-amino acids (**9**) and (**23**) (see *figs. 79 & 80*). Peaks at 4.1 ppm and 5.3 ppm corresponding to Fmoc H-10 and H-9 were also observed.

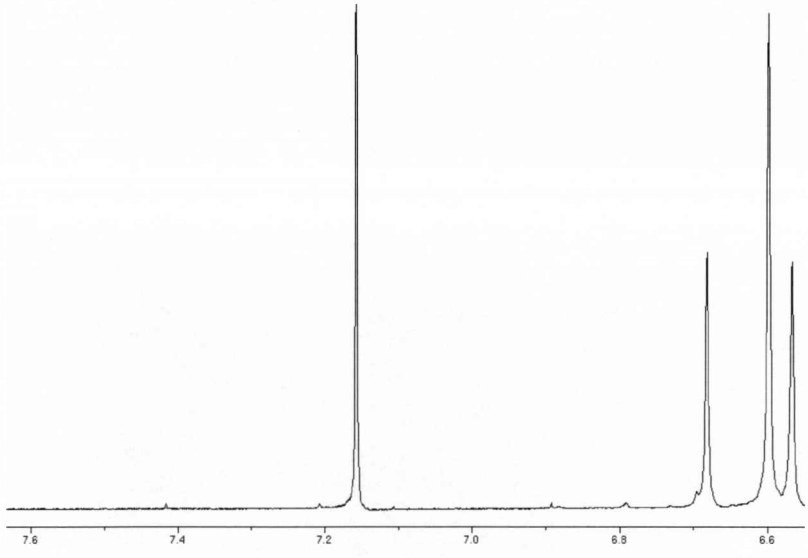


Figure 79: *Aromatic region of 1D NMR of FD-O(H).*

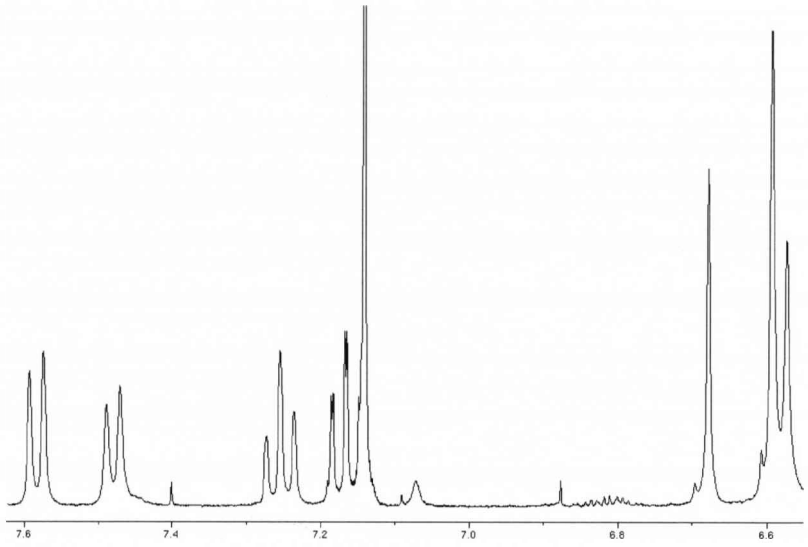


***Figure 80**: Aromatic region of 1D NMR of FD-O-Gly-NHFmoc (**33**)*

The peak of the glycine CH_2 was also plainly visible at 4.4 ppm. As for any solution-phase conversion it was also possible to establish an isolated yield (88 %) for the glycine derived dendrimer; a measurement which is unattainable in SPPS without termination of the synthesis. Identification of the glycine derived dendrimer (**33**) was later confirmed by MALDI-TOF mass spectrometry.

In order to test if the SPPS conditions used in previous syntheses of nucleoside β-peptides were directly transferrable to the fluorous-phase procedure, deprotection of the Fmoc group was attempted with a biphasic mixture of FC-72 and 2 % DBU/DMF. The Fmoc group was successfully removed within a few minutes of exposure to the base, however, significant cleavage of the glycine from the dendrimer was also observed. This is most likely due to formation of an oxazolone derivative by attack of the imidate oxygen generated by deprotonation of the Fmoc amide. This mechanism is similar to racemisation of activated N-acyl protected amino acids (see *fig. 81*). This problem is commonly encountered in PNA synthesis[122] and in this case resulted in the cleaved dendrimer being the major product isolated in the fluorous layer after work-up. To suppress this side reaction the reactive imidate could be trapped out by the addition of a silylating agent to form a silyl enol ether and this could subsequently be hydrolysed by an acidic wash which would also serve to neutralise the DBU.

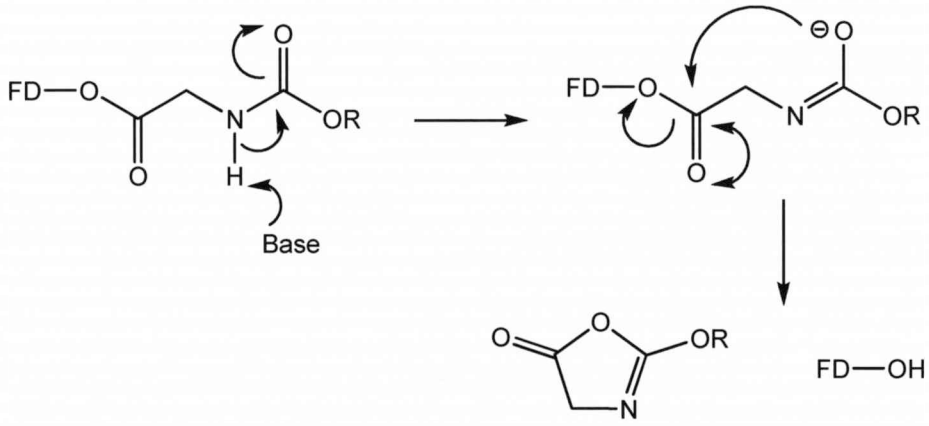


Figure 81: *N-acyl transfer resulting in cleavage of glycine from the fluorous dendrimer.*

A more straightforward approach was to replace DBU with the less basic piperidine which has been observed not to give acyl migration products during Fmoc synthesis of PNA.[228] Successful deprotection was achieved by treatment with a 10 % solution of piperidine in DMF in biphasic conditions for 30 minutes. The NMR spectrum showed disappearance of the Fmoc resonances and shift of the glycine and dendrimer methylene protons to 3.56 and 5.18 ppm respectively.

Thymidine nucleoside amino acid (**9**) was condensed with the glycine derived dendrimer using HBTU in a THF suspension in the presence of 2,4,6-collidine. After 24 hours at ambient temperature the reaction was not complete; the product to starting material ratio was approximately 3:1 by NMR. Peaks corresponding to H1', H3' and H4' of the sugar ring were all clearly visible alongside the same aromatic pattern for the Fmoc protons as seen with glycine. The reaction proceeded no further after extended periods of time. Heating the reaction to 45 °C also failed to drive the reaction beyond approximately 50 % completion. It is thought that the major barrier to a smooth reaction is the vastly different solubility of the dendron compared to the amino acid. Whilst THF is a suitable solvent for the dendron, the nucleoside was only sparingly soluble and caused aggregation of the fluorous component when coupled. Although encouraging coupling results were achieved in this solvent, the system requires further optimisation.

Cleavage of the dipeptide unit was achieved by treatment of a THF solution of the dendrimer with concentrated ammonia at ambient temperature for 2 hours. After partition with FC-72 the fluorous dendron was recovered in 93 % yield

demonstrating that recycling of the fluorous support would be possible after the completion of a synthetic cycle.

3.6 Limitations and Extensions of FPS for Peptide Synthesis

One of the more challenging aspects of the use of heavy fluorous chemistry for oligomerisation protocols is the need to maintain the ratio of 60 % fluorine for efficient partition in extractive work-up. In the case of the thymidine β amino acid (**9**) coupling only two residues, glycine and a single thymidine amino acid, to the fluorous dendrimer reduces the fluorine ratio to under 56 %. As the peptide chain is lengthened, the fluorine content will progressively diminish until it reaches a level at which partition with fluorous solvents is no longer effective.

This problem could potentially be solved by the inclusion of a fluorous moiety in the amino acids themselves. The best prospect for introduction of such a group would be as part of the protecting strategy. Fluorous fmoc (see *fig. 82*) is commercially available but presents only a limited opportunity to improve the fluorine ratio of any peptide if it is only incorporated as the *N*-terminal protecting group. The other site that requires protection in this type of amino acid is the nucleobase. Whilst a protecting group for thymine was not explicitly necessary during the synthesis or oligomerisation of the T amino acid, the exocyclic amine of adenine required masking to prevent unwanted side reactions from occurring. Protection of the nucleobase with a fluorous group could provide the additional fluorines required for successful peptide synthesis by heavy fluorous methods.

It is also interesting to consider the use of fluorous protecting groups for conventional SPPS. Benzoyl protection of the exocyclic amine of the dA derived amino acid was found to be problematic in terms of deprotection and an alternative is required for production of fully deprotected peptides. Orthogonal, acid-labile protecting groups are an attractive option; however, when removed at the same time as the peptide is cleaved from the resin, these groups present an obvious impurity in the final product. These impurities would require removal by HPLC and this would impose limitations on the scale of production. If it were possible to introduce a fluorous, acid-sensitive functionality as the nucleobase protecting group, for example a fluorous trityl derivative, then it is, in principle, possible to consider large-scale, fluorous-based purification methods as an alternative to HPLC.

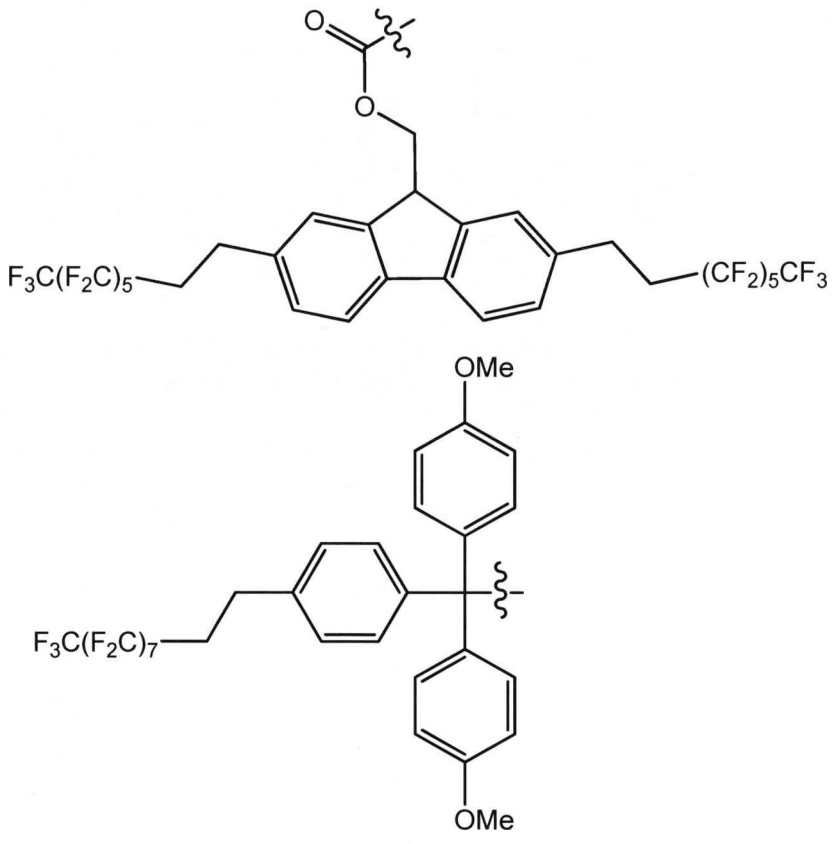


Figure 82: *Fluorous fmoc (top) and dimethoxytrityl (bottom).*

The fluorous trityl derivatives which are already commercially available (see *fig. 80*)[273] would not be appropriate for heavy fluorous chemistry, as the fluorine ratio is too low, but would be ideally suited to a standard F-SPE process. Conversely, the use of fluorous Fmoc on the peptide *N*-terminus and an organic nucleobase protecting group would also have the same potential for separation by standard F-SPE and this type of procedure has already been successfully demonstrated with small dipeptides[276].

3.7 Conclusion

A model peptide synthesis using fluorous-phase chemistry with nucleoside β-amino acids has been performed. The intermediate compounds from each of the individual deprotection, coupling and cleavage steps were all purified without the need for chromatography and analysed whilst still bound to the fluorous support. Equivalent purification and analysis is not possible with traditional SPPS approaches.

FPS is, however, not without issues which will require addressing for the future of this technology. The maintenance of the 60 % fluorine ratio has already been discussed in the previous section and this is something which can be considered in conjunction with a reaction that is yet to be attempted by fluorous means; capping of failure sequences. Capping could possibly be attempted in either single phase system in a manner similar to the coupling step or a biphasic approach akin to the deprotection step. It is not easy to speculate on which will offer the most efficient pathway, as much will depend on the solvent compatibility of the capping reagent with the fluorous or organic solvent systems. Optimum conditions should be relatively straightforward to determine experimentally. Failure sequences are also as much of a problem in this type of FPPS as in SPPS. A possible solution is to, once more, utilise the 60 % ratio as a handle for separating complete from deletion sequences by differentially capping the completed peptide. This, however, would be far from ideal as the capping reagent would require a large molecular weight in order to significantly alter the fluorine ratio. Even if such a cap could successfully be attached to the peptide, efficient partition would not be guaranteed.

There are many more possible and alternative opportunities to utilise FPS chemistry in peptide synthesis. This procedure most of all represents one of the major attractive aspects; the ease of purification of fluorous compounds. Isolation of oligonucleotides and peptides made by more conventional means often relies on laborious and time-consuming HPLC and the potential for eliminating this has been clearly demonstrated.

Chapter 4

Experimental

4 Experimental

Techniques and Reagents

Analytical thin layer chromatography (tlc)

Performed on 2 cm x 5 cm Merck silica gel 60 F_{254} aluminium plates. Plates were visualised with a 254 nm UV lamp before staining with *p*-anisaldehyde "dip" (see general reagents). Nucleoside compounds were stained blue on exposure to a heat gun.

Flash column chromatography

The required quantity of silica gel (40 – 63 μm particle size from BDH) was made into a slurry with column eluent and applied to the column over a sand base layer. Crude material was introduced in either a minimal volume of column eluent or adsorbed onto silica.

High Performance Liquid Chromatography (HPLC)

HPLC was performed on an automated Gilson HPLC system equipped with an autoinjector, a photodiode array detector and a dual hydraulic pump. Chromatographic data was handled using UniPoint Version 3.0 and absorptions measured at 254 nm.

Separations were performed on a Phenomenex® reverse-phase 250 mm x 4.6 mm, Gemini 5u C18 110 Å column using a gradient of 0 – 30 % MeCN in aqueous triethylammonium bicarbonate (TEAB, 0.1 M) for thymidine oligomers and 0 – 40 % MeCN in aqueous triethylammonium bicarbonate (TEAB, 0.1 M) for deoxyadenosine oligomer over 30 minutes. The gradient was followed by a 10 minute column flush with the 30 or 40 % MeCN eluent.

NMR Spectroscopy

Spectra were recorded on a Bruker AMX-400 MHz machine. Coupling constants are reported in Hz and chemical shift values are given in ppm from an internal standard of tetramethylsilane (TMS). Chemical shift values for some aromatic protons, fluorenyl protons and carbons have been assigned using simulations from ChemDraw 10™.

IR Spectroscopy

Spectra for thymidine compounds were recorded in nujol mulls on a Perkin Elmer FT-IR Paragon 1000 spectrometer. 2'-Deoxyadenosine samples were recorded neat on a Jasco FT-IR 4100 spectrometer equipped with an ATR sample holder.

UV Spectroscopy

Spectra were recorded at the stated wavelengths on a Hewlett Packard 8452A spectrodiode array in 1 cm quartz cuvettes.

Mass Spectrometry

Electrospray spectra were recorded by Mr Allan Mills of the University of Liverpool mass spectrometry service on a Micromass LCT using electrospray ionisation (EI) in the specified ionisation mode with a direct infusion syringe pump.

CHN Microanalysis

Obtained by Mr Steve Apter of the University of Liverpool microanalysis service on a Thermo FlashEA 1112 series CHNSO analyser.

Melting Point

Where physical form allowed, melting point was determined on a Gallenkamp melting point apparatus.

Solvents

Unless stated solvents were purchased from Fisher Scientific or BDH. Anhydrous solvents were obtained as follows:

Dichloromethane: Distilled from calcium hydride.

N,N-Dimethylacetamide, *N,N*-dimethylformamide & 1,4-dioxane: Sure-seal™ anhydrous reagents purchased from Sigma-Aldrich.

Methanol: Distilled from magnesium and iodine.

Pyridine: Distilled from calcium hydride

General Reagents

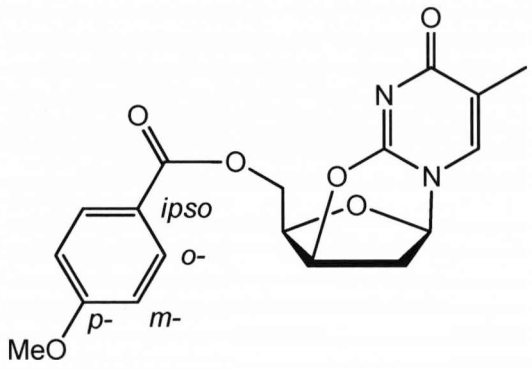
Unless stated, reagents were purchased from Sigma-Aldrich, Fluka or BDH and used as supplied. Thymidine was purchased from Alfa Aesar, 2'-deoxyadenosine

was purchased from ChemGenes (USA) in monohydrate form and dried by co-evaporation with anhydrous pyridine before use. 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HBTU) and benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP®) were purchased from Lancaster Synthesis and stored at 4 °C over silica desiccant before use..

p-Anisaldehyde "dip" was prepared as follows: *p*-Anisaldehyde (6 ml) was mixed with sulfuric acid (8.0 ml), acetic acid (2.4 ml) and ethanol (218 ml) and stored at room temperature in darkness.

4.1 Synthesis of 3'-N-(9-fluoromethoxycarbonyl)-amino-3'-deoxythymidine-5'-carboxylic acid

Preparation of 2-3'-anhydro-5'-O-(4-*O*-methoxybenzoyl) thymidine (5)



Thymidine (10.0 g, 41 mmol) and PPh_3 (16.2 g, 62 mmol) were dissolved in anhydrous DMF (120 ml). To this stirred mixture was added *p*-methoxybenzoic acid (9.4 g, 62 mmol) followed by DIAD (12.2 ml, 62 mmol) drop-wise over 30 minutes. The reaction vessel was kept at < 2 °C in an ice/water bath during this addition. The solution was left to stir at ambient temperature for 15 minutes then a further aliquot of PPh_3 (19.6 g, 75 mmol) and DIAD (14.7 ml, 75 mmol) was added in the same manner. After half the DIAD addition, an off-white precipitate had appeared. The reaction mixture was left to stir at ambient temperature for a further 2 hours then poured into Et_2O (1250 ml) and cooled to 4 °C overnight. The white powdery precipitate was filtered off and thoroughly washed with Et_2O.

Yield: 14.9 g, 41 mmol, > 99 %, m.pt. > 250 °C

δ^1H (400 MHz, DMSO-d_6) 1.74 (3H, s, CH_3, Thy), 2.52 (1H, m, H2'), 2.61 (1H, m, H2'), (3H, s, CH_3O), 4.31 (1H, dd, $^2J_{5'}$ = 11.76 Hz, $^3J_{4'}$ = 6.44 Hz, H5'), 4.48 (1H, dd, $^2J_{5'}$ = 11.84 Hz, $^3J_{4'}$ = 5.20 Hz, H5'), 4.57 (1H, dt, $^3J_{5'}$ = 5.74 Hz, $^3J_{3'}$ = 2.36 Hz, H4'), 5.41 (1H, bs, H3'), 5.89 (1H, pseudo-d, J = 3.80 Hz, H1'), 7.00 (2H, d, J = 7.28 Hz, *o*-Ar-H), 7.56 (1H, d, $^4J_{CH3}$ = 0.96 Hz, H-6), 7.86 (2H, d, J = 8.92 Hz, *m*-Ar-H),

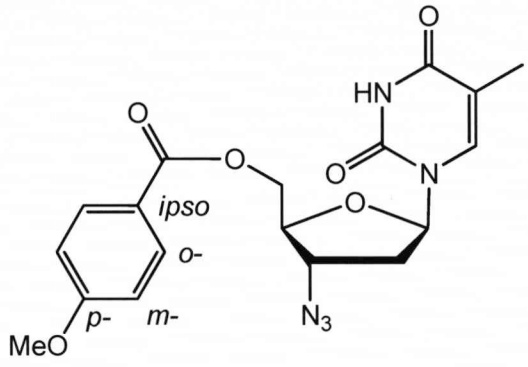
δ^{13}C (100 MHz, DMSO-d_6) 13.3 (CH_3, Thy), 33.1 (C2'), 55.9 (CH_3O), 62.6 (C5'), 77.6 (C3'), 82.4 (C4'), 87.3 (C1'), 114.4 (*m*-Ar-C), 116.5 (C5), 121.7 (ipso-Ar-C),

131.7 (*o*-Ar-C), 137.0 (C6), 153.7 (C2), 163.7 (*p*-Ar-C), 165.4 (COO, anisyl), 171.2 (C4).

Analysis found C 59.97 H 5.00 N 7.82; $C_{18}H_{18}N_2O_6$ requires C 60.30 H 5.02 N 7.82

Found m/z (ES+) 381.1073 ([M + Na]$^+$ 100 %); [$C_{18}H_{18}N_2O_6$ + Na]$^+$ requires 381.1063

Preparation of 3'-azido-3'-deoxy-5'-*O*-(4-methoxybenzoyl)-thymidine (6)



2'-3'-Anhydro-5'-O-(4-*O*-methoxybenzoyl) thymidine (**5**, 11.7 g, 32.5 mmol) and LiN_3 (4.8 g, 98 mmol) were dissolved in anhydrous DMF (90 ml) and heated to 125 °C for 4 hours using an oil bath. After this period no starting material remained and the orange coloured reaction mixture was poured into 5% aqueous HCl (295 ml) then extracted with EtOAc (3 x 600 ml). The combined organic extracts were washed with water (500 ml) and brine (500 ml) then dried over $MgSO_4$. The solvents were reduced to a minimum volume under vacuum and the crude product was purified by flash column (EtOAc : hexane, 2 : 1). The pure product was obtained as a white foam after evaporation of the chromatography solvents.

Yield: 8.6 g, 24.7 mmol, 65 %

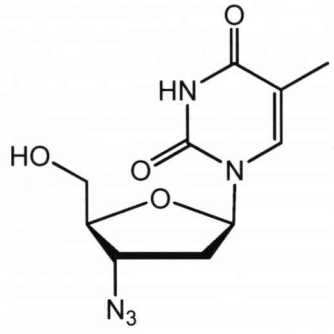
δ^1H (400 MHz, $CDCl_3$) 1.72 (3H, s, CH_3, Thy), 2.36 (1H, m, H2') 2.54 (1H, m, H2'), (3H, s, CH_3O), 4.21 (1H, pseudo-q, H4'), 4.36 (1H, dt, ^3J = 7.4 Hz, ^3J = 4.89 Hz, H3'), 4.53, (1H, dd, $^2J_{5'}$ = 12.32 Hz, $^3J_{4'}$ = 4.00 Hz, H5'), 4.64 (1H, dd, $^2J_{5'}$ = 12.44 Hz, $^3J_{4'}$ = 3.72 Hz, H5'), 6.17 (1H, t, $^3J_{2'}$ = 6.24 Hz, H1'), 6.93 (2H, m, *o*-Ar-H) 7.22 (1H, d, $^4J_{CH3}$ = 1.12 Hz, H-6), 7.98 (2H, *m*-Ar-H), 9.28 (br s, NH).

$\delta^{13}C$ (100 MHz, $CDCl_3$) 12.6 (CH_3, Thy), 38.2 (C2'), 55.9 (CH_3O), 61.1 (C3'), 63.8 (C5'), 82.6 (C4'), 85.8 (C1'), 111.7 (C5), 114.4 (2 x m-Ar-C), 121.9 (ipso-Ar-C), 132.3 (2 x o-Ar-C), 135.4 (C6), 150.6 (C2), 164.3 (C4), 164.4(p-Ar-C), 166.1 (anisyl COO).

Found m/z (ES+) 424.1231 ($[M + Na]^+$ 100 %); $[C_{18}H_{18}N_2O_6 + Na]^+$ requires 424.1233

IR Nujol mull: ν = 2099.9 cm^{-1} (azide asymmetric stretch)

Preparation of 3'-azido-3'-deoxythymidine (2)



3'-Azido-3'-deoxy-5'-*O*-(4-methoxybenzoyl) thymidine (**6**, 6 g, 15 mmol) was dissolved in MeOH (60 ml) and NaOMe (3.24 g, 60 mmol) in H_2O solution (66 ml) was added and the reaction stirred for 3.5 hours. Upon further addition of H_2O (60 ml) white crystals, later analysed to be *p*-methoxybenzoic acid, were formed and isolated by filtration. After a further crystallisation with H_2O (60 ml) the aqueous solution was washed with Et_2O (2 x 60 ml) and Amberlite® H^+ ion exchange resin (12 g) added. The mixture was stirred for 15 minutes then the Amberlite removed by filtration and the solvents removed under reduced pressure to give a yellow glassy solid. The crude product was purified by flash column (EtOAc : MeOH, 4:1) to give AZT as a fine off-white powder after removal of the chromatography solvents.

Yield: 3.14 g, 11.8 mmol, 78 %, m.pt. 110-112 °C

δ^1H (400 MHz, MeOH-d_4) (1.88, 3H, s, CH_3, Thy), 2.40 (2H, m, H2'), 3.74 (1H, dd, $^2J_{5'}$ = 12.14 Hz, $^3J_{4'}$ = 3.22 Hz, H5') 3.83 (1H, dd, $^2J_{5'}$ = 12.14 Hz, $^3J_{4'}$ = 3.42 Hz, H5'),

3.91 (1H, dt, $^3J_{3'}$ = 4.81 Hz, $^3J_{5'}$ = 3.24 Hz, H4'), 4.35 (1H, m, H3'), 6.17 (1H, t, $^3J_{2'}$ = 6.44 Hz, H1'), 7.78 (1H, s, H-6).

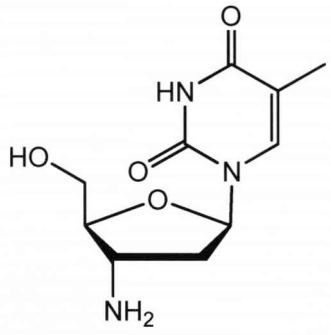
δ^{13}C (100 MHz, MeOH-d_4) 12.8 (CH_3, Thy), 38.7 (C2'), 62.0 (C5'), 62.9 (C3'), 86.5 (x2, C1', C4'), 112.0 (C5), 138.4 (C6), 152.7 (C2), 166.8 (C4).

Analysis found C 44.98 H 4.88 N 26.02; $C_{10}H_{13}N_5O_4$ requires C 44.98 H 4.91 N 26.22

Found m/z (ES+) 290.0876 ([M + Na]$^+$ 100 %); [$C_{10}H_{13}N_5O_4$ + Na]$^+$ requires 290.0865

IR Nujol mull: ν = 2081.7 cm^{-1} (azide asymmetric stretch)

Preparation of 3'-amino-3'-deoxythymidine (7)



3'-Azido-3'-deoxythymidine (**2**, 2.7 g, 10.0 mmol) was dissolved in anhydrous MeOH (27 ml) and palladium on activated carbon (1 microspatula) added to the vigorously stirred solution. The reaction vessel was then evacuated and flushed with H_2 (g) three times then stirred for 4 hours. The Pd catalyst was removed by filtration through Celite® and the filter medium washed thoroughly with MeOH. The solvent was removed under reduced pressure to give the product as an amorphous white foam.

Yield: 2.4 g, 9.9 mmol, 99 %

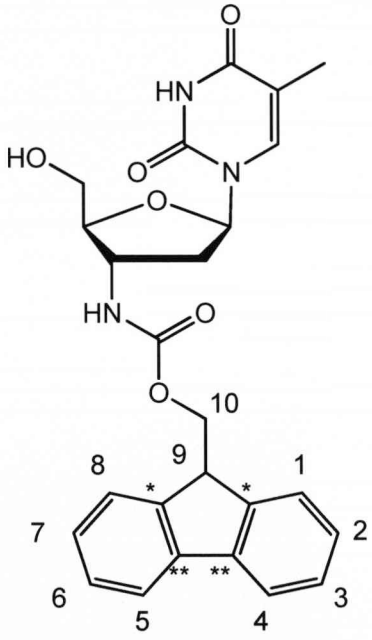
δ^1H (400 MHz, MeOH-d_4) 1.88 (3H, s, CH_3 Thy), 2.21 (1H, m, H2'), 2.28 (1H, m, H2'), 3.56 (1H, pseudo-q, H3'), 3.73 (1H, m, H4'), 3.80 (1H, m, H5') 3.86 (1H, dd,

$^{2}J_{5'}$ = 12.16 Hz, $^{3}J_{4'}$ = 3.04 Hz, H5'), 6.18 (1H, dd, ^{3}J = 6.94 Hz, ^{3}J = 4.82 Hz, H1'), 7.85 (1H, d, $^{4}J_{CH3}$ = 1.12 Hz, H-6).

δ^{13}C (100 MHz, MeOH-d_4) 12.9 ($\underline{C}H_3$, Thy), 42.1 (C2'), 50.1 (C3'), 62.6 (C5'), 86.4 (C1'), 89.2 (C4'), 111.7 (C5), 138.7 (C6), 152.9 (C2), 167.0 (C4).

Found m/z (ES+) 264.0960 ([M + Na]$^+$ 100 %); [$C_{10}H_{15}N_3O_4$ + Na]$^+$ requires 264.0996, 527.1842 ([2M – H + 2Na]$^+$ 30.60 %); [$C_{20}H_{29}N_6O_8$ + 2Na]$^+$ requires 527.1846

Preparation of 3'-*N*-(9-fluoromethoxycarbonyl)-3'-deoxythymidine (8)



3'-Amino-3'-deoxythymidine (**7**, 190 mg, 3.7 mmol) was dissolved in 1,4-dioxane : water (2 : 1, 4 ml) and K_2CO_3 (218 mg, 1.6 mmol) added. The reaction mixture was allowed to stir for 30 minutes at ambient temperature before being reduced to 0 °C in an ice bath. FmocCl (265 mg, 1.0 mmol) was added and product as a white powdery solid was observed to precipitate out of solution almost immediately. After 30 minutes stirring at ambient temperature the solution had turned thick and opaque, therefore water (30 ml) was added, causing more white solid product to precipitate out which was collected by filtration and washed thoroughly with water.

Yield: 276 mg, 0.6 mmol, 76 %, m.pt.178-181 °C

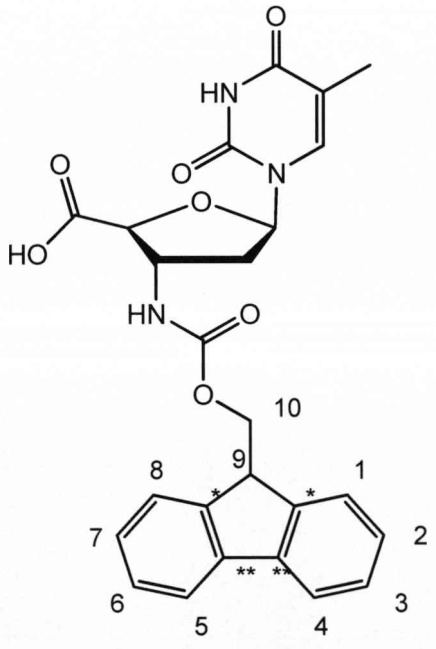
δ^1H (400 MHz, MeOH-d_4) 1.87 (3H, s, CH_3 Thy), 2.29 (2H, m, H2'), 3.68 (1H, m, H4'), 3.82 (2H, m, H5'), 4.19 (1H, t, $^3J_{10}$ = 6.28 Hz, Fmoc H-9), 4.26 (1H, pseudo-q, H3') 4.40 (2H, m, Fmoc H-10) 6.18 (1H, t, $^3J_{2'}$ = 5.98 Hz, H1'), 7.29, (2H, m, Fmoc H-2, 7), 7.33, (2H, m, Fmoc H-3, 6) 7.63 (2H, d, $^3J_{2,7}$ = 7.00 Hz, Fmoc H-1, 8), 7.78 (2H, d, $^3J_{3,6}$ = 7.40 Hz, Fmoc H-4, 5), 7.86 (1H, s, H-6).

δ^{13}C (100 MHz, MeOH-d_4) 12.5 ($\underline{C}H_3$, Thy), 39.0 (C2'), 49.0 (Fmoc C-9) 51.6 (C3'), 62.3 (C5'), 67.7 (Fmoc C-10) 85.8 (C1'), 86.5 (C4'), 111.5 (C5), 121.0 (Fmoc C-4, 5), 126.1 (Fmoc C-1, 8), 128.2 (Fmoc C-2, 7), 128.8 (Fmoc C-3, 6), 138.2 (C6), 142.7 (Fmoc-C**), 145.2 (Fmoc-C*), 152.3 (C2), 158.4 (Fmoc C=O) 166.5 (C4).

Analysis found C 64.60 H 5.59 N 8.89; $C_{25}H_{25}N_3O_6$ requires C 64.79 H 5.44 N 9.07

Found m/z (ES+) 486.1637 ([M + Na]$^+$ 100%); [$C_{25}H_{25}N_3O_6$ + Na]$^+$ requires 486.1641

Preparation of 3'-*N*-(9-fluoromethoxycarbonyl)-3'-deoxythymidyine-5'-carboxylic acid (1)



3'-N-(9-fluoromethoxycarbonyl)-3'-deoxythymidine (**8**, 5.4 g, 11.6 mmol) was suspended in MeCN/H_2O (108 ml). To this suspension TEMPO (1.8 g, 11.6 mmol)

and BAIB (8.2 g, 25.6 mmol) were added and the bright orange mixture stirred at ambient for 3 hours. After this time product in the form of a pale brown precipitate had formed and was isolated by filtration and washed with Et_2O. After a further 12 hours stirring at ambient, a further crop of pale brown precipitate had formed and the solution had separated into two distinct layers. The precipitate was isolated by filtration and washed with Et_2O. The remaining reaction mixture was left for a further 24 hours at ambient and a further crop of precipitate isolated and washed.

Yield: 1.98 g, 4.1 mmol, 46 %

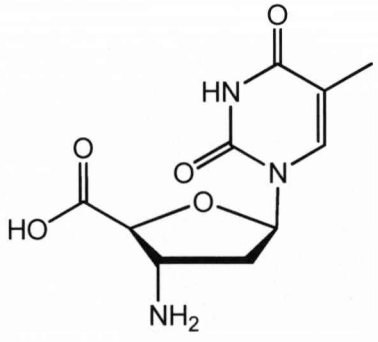
δ^1H (400 MHz, MeOH-d_4) 1.80 (3H, s, CH_3 Thy), 2.25 (2H, bs, H2'), 4.25 (1H, t, $^3J_{10}$= 6.64 Hz, Fmoc H-9), 4.30 (1H, d, $^3J_{3'}$ = 3.76 Hz, H4'), 4.35 (3H, m, H-10, H3'), 6.34 (1H, m, H1'), 7.35 (2H, m, Fmoc H-2, 7), 7.42 (2H, m, Fmoc H-3, 6), 7.70 (2H, d, $^3J_{2,7}$ = 7.04 Hz, Fmoc H-1, 8), 7.89 (2H, d, $^3J_{3,6}$ = 7.40 Hz, Fmoc H-4, 5), 7.98 (1H, s, H-6), 8.04, (1H, bs, 3'-NH), 11.31 (bs, NH).

δ^{13}C (100 MHz, MeOH-d_4) 12.7 (CH_3, Thy), 36.5 (C2'), 47.1 (Fmoc C-9), 54.8 (C3') 66.0 (Fmoc CH_2) 81.8 (C4'), 85.5 (C1'), 109.7 (C5), 120.1(Fmoc C-4, 5), 125.5 (Fmoc C-1, 8), 127.4 (Fmoc C-2, 7), 128.0 (Fmoc C-3, 6), 136.7 (C6), 141.1 (Fmoc-C**), 144.1 (Fmoc-C*), 150.8 (C2), 155.9 (Fmoc C=O) 164.0 (C4), 172.5 (C5').

Analysis found C 62.60 H 4.78 N 8.68; $C_{25}H_{23}N_3O_7$ requires C 62.89 H 4.85 N 8.80

Found m/z (ES^-) 476.1466 ([M - H]$^-$ 100%); [$C_{25}H_{23}N_3O_7$ - H]$^-$ requires 476.1548

Preparation of 3'-amino-3'-deoxythymidine-5'-carboxylic acid (9)



To a stirred solution of 3'-amino-3'-deoxythymidine (**2,** 2.4 g, 9.9 mmol) in MeCN/ H_2O (1 : 1, 72 ml) was added TEMPO (310 mg, 2 mmol) and BAIB (7 g, 21.0 mmol) and the reaction stirred at ambient temperature for 4 hours. At this point a white precipitate of product had formed in the upper layer of the now biphasic orange mixture which was isolated by filtration. The remaining solution was stirred for a further 24 hours then further TEMPO (1.2 g, 8.0 mmol) and BAIB (3.2 g, 10.5 mmol) added. After 30 minutes further stirring more white precipitate was formed and isolated by filtration, which, after drying, was a very fine powder-like solid.

Yield: 592 mg, 2.3 mmol, 23 %

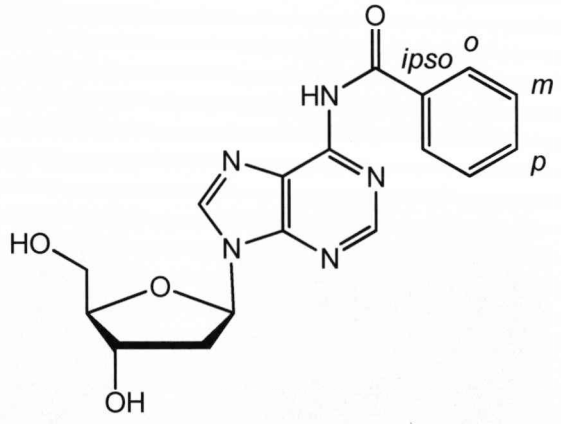
δ^1H (400 MHz, DMSO-d_6) 2.15 (1H, m, H2'), 2.29 (1H, m, H2''), 3.86 (1H, m, H3'), 4.19 (1H, d, $^3J_{3'}$ = 2.84 Hz, H4'), 6.32 (1H, t, $^3J_{2'}$ = 6.64 Hz, H1'), 8.75 (1H, s, H-6).

δ^{13}C (100 MHz, DMSO-d_6) 12.7 (CH_3, Thy), 36.6 (C2'), 54.7 (C3'), 83.8 (C4'), 85.3 (C1'), 109.1 (C5), 138.1 (C6), 150.9 (C2), 164.3 (C4) 172.6 (C5').

Found m/z (ES^-) 254.0773 ([M - H]$^-$ 100%); [$C_{10}H_{12}N_3O_5$ - H]$^-$ requires 254.0777

4.2 Synthesis of N6-benzoyl-3'-N-(9-fluorenylmethoxycarbonyl)-amino-2'-3'-dideoxyadenosine-5'-carboxylic acid

Preparation of *N*6-benzoyl-2'-deoxyadenosine (16)



2'-Deoxyadenosine monohydrate (9.1 g, 34 mmol) was dried by co-evaporation with dry pyridine (2 x 50 ml) and suspended in dry pyridine (185 ml). To this suspension

was added dry Et_3N (14.2 ml, 102 mmol) followed 30 minutes later by TMSCl (21.6 ml, 170 mmol). A further aliquot of dry pyridine (40 ml) was added to facilitate stirring of the now thick pink slurry. After a further 30 minutes stirring BzCl (5.9 ml, 51 mmol) was added and the slurry left to stir for 2 hours. Ice bath cooling was applied to the reaction vessel before H_2O (45 ml) and cNH_3 (45 ml) were added and the now yellow solution stirred at ambient temperature for 30 minutes. The resulting orange solution was evaporated under reduced pressure, dissolved in H_2O (100 ml) and washed with EtOAc (100 ml). Crystallisation of the product began in the separating funnel and a thick slurry was obtained by the time separation was complete. The white solid was filtered off and the remaining solution refrigerated overnight to yield a second crop of product as a powdery white solid.

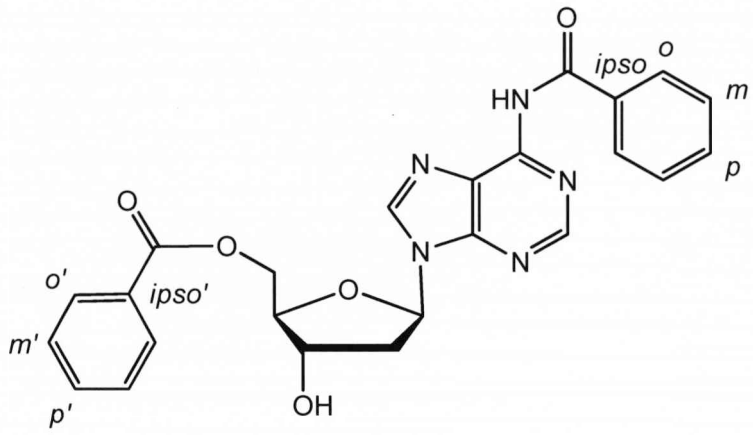
Yield: 9.0 g, 25.3 mmol, 75 %

δ^1H (400 MHz, DMSO-d_6) 2.38 (1H, ddd, $^2J_{2'}$ = 13.28 Hz, $^3J_{1'}$ = 6.24Hz, $^3J_{3'}$ = 3.44 Hz, H2'), 2.81 (1H, m, H2'), 3.55 (1H, dd, $^2J_{5'}$ = 11.66 Hz, $^3J_{4'}$ = 4.46, H5''), 3.64 (1H, dd, $^2J_{5'}$ = 11.76 Hz, $^3J_{4'}$ = 4.56 Hz, H5'), 3.91 (1H, pseudo-q, H4'), 4.47 (1H, m, H3'), 6.49 (1H, t, $^3J_{2'}$ = 6.74 Hz, H1'), 7.55 (2H, t, ^3J_p = 7.5 Hz, *m*-Ar-H), 7.65 (1H, t, ^3J_m = 7.40 Hz, *p*-Ar-H), 8.05 (2H, d, ^3J_m = 7.20 Hz, *o*-Ar-H), 8.67 (1H, s, H8), 8.75 (1H, s, H2), 11.16 (bs, NH).

δ^{13}C (100 MHz, DMSO-d_6) 39.7 (C2'), 62.0 (C5'), 71.1 (C3'), 84.1 (C1'), 88.4 (C4'), 126.3 (C5), 128.8 (*o*-, *m*-Ar-C), 132.8 (*p*-Ar-C), 133.8 (*ipso*-Ar-C), 143.4 (C8), 150.7 (C6), 151.9 (C2), 152.3 (C4), 166.0 (C=O).

Found m/z (ES^+) 378.1161 ($[M + Na]^+$ 100%); $[C_{17}H_{17}N_5O_4 + Na]^+$ requires 378.1178

Preparation of *N*-6, 5'-*O*-dibenzoyl-2'-deoxyadenosine (17)



2'-Deoxyadenosine (2 g, 8.0 mmol) was dried by co-evaporation with dry pyridine (2 x 10 ml) then dissolved in dry pyridine (30 ml). BzCl (1 ml, 8.8 mmol) in dry pyridine solution (20 ml) was added dropwise over approximately 30 minutes and after 15 minutes another portion BzCl (0.2 ml 1.8 mmol) was added and the reaction stirred for a further 30 minutes. To the now bright green solution TMSCl (5.1 ml, 40 mmol) and BzCl (2.8 ml, 24 mmol) were added in succession, dropwise, and the cloudy green mixture stirred for 1.5 hours. With the reaction vessel in ice bath cooling a MeOH (10 ml) quench was applied before cNH_3 (10 ml) was dropped in and the green slurry stirred at 0 °C for 30 minutes. The suspension was evaporated under reduced pressure before being partitioned between $CHCl_3$ and H_2O (50 ml each). The organic layer was separated, dried over $MgSO_4$ and evaporated to give a dark green oil which was purified by flash column (5 % MeOH/ $CHCl_3$) to give the desired product as a white foam after solvent removal.

Yield: 2.9 g, 6.4 mmol, 80 %

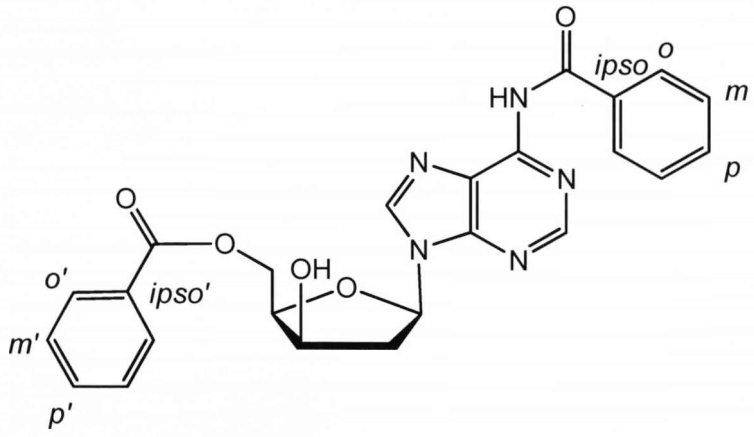
δ^1H (400 MHz, DMSO-d_6) 2.43 (1H, m, H2''), 2.96 (1H, m, H2'), 4.15 (1H, dt, $^3J_{5'}$ = 5.72 Hz, $^3J_{3'}$ = 4.28, H4'), 4.41 (1H, dd, $^2J_{5'}$ = 11.76 Hz, $^3J_{4'}$ = 5.88, H5'), 4.52 (1H, dd, $^2J_{5'}$ = 11.86 Hz, $^2J_{4'}$ = 4.26 Hz, H5'), 4.64 (1H, bs, H3'), 5.53 (1H, bs, 3'-OH), 6.47 (1H, t, $^3J_{2'}$ = 6.64 Hz, H1'), 7.45 – 7.51 (4H, m, *m*-, *m'*-Ar-H), 7.57 (2H, m, *p*-, *p'*-Ar-H), 7.86, 7.99 (4H, m, *o*-, *o'*- Ar-H), 8.58 (1H, s, H8), 8.65 (1H, s, H2), 11.09 (bs, NH).

$\delta^{13}C$ (100 MHz, DMSO-d_6) 38.8 (C2'), 64.7 (C5'), 70.9 (C3'), 84.1 (C1'), 84.6 (C4'), 126.3 (C5), 127.8 – 129.5, 131.5 – 133.8 (10 x Ar-C), 129.8 (*ipso'*-Ar-C), 134.7 (*ipso*-Ar-C), 143.7 (C8), 150.8 (C6), 151.9 (C2), 152.2 (C4), 165.9 (5'-OBz C=O), 168.3 (NHBz C=O).

Found m/z (ES^+) 460.1630 ($[M + H]^+$ 100%); $[C_{24}H_{22}N_5O_5]^+$ requires 460.1621

Analysis found C 62.72 H 4.61 N 15.19; $C_{24}H_{21}N_5O_5$ requires C 62.74 H 4.61 N 15.24

Preparation of *N*-6, 5'-*O*-dibenzoyl-3'-*xylo*-2'-deoxyadenosine (18) (and *N*-6, 3'-*O*-dibenzoyl-3'-*xylo*-2'-deoxyadenosine (22))



Method A:

N-6-5'-Dibenzoyl-2'-deoxyadenosine (**17**, 5.4 g, 11.8 mmol) was dissolved in DCM (160 ml) containing pyridine (2.8 ml, 35.4 mmol) and the solution cooled to -30 °C in a dry ice/EtOAc bath. To the cold solution was added $(TfO)_2O$ (2.8 ml, 35.4 mmol, 10 % v/v solution in DCM) dropwise. Following the addition the temperature of the reaction was raised to 0 °C over 30 minutes and $PhCO_2H$ (4.3 g, 35.4 mmol) and H_2O (11.7 ml) added. The solution was stirred overnight at approximately 0 °C then the organic layer separated, dried over MgSO4 and evaporated. The crude mixture of isomers was purified by flash column (0 – 3 % MeOH/$CHCl_3$) to give both products as pale yellow foams after solvent removal.

Yield: 1.23 g, 2.7 mmol, 23 %

Method B:

N-6-5'-Dibenzoyl-2'-deoxyadenosine (**17**, 5.4 g, 11.8 mmol) was dissolved in DCM (250 ml) containing pyridine (10 ml) and the solution cooled to -30 °C in a dry ice/EtOAc bath. $(TfO)_2O$ (2.8 ml, 35.4 mmol, 10 % v/v solution in DCM) was introduced dropwise then the cooling bath removed and the reaction stirred for a further 45 minutes before water (5 ml) was poured in. The resulting emulsion was stirred at ambient for 2 hours then a further aliquot of water (25 ml) added and the organic layer separated, dried over $MgSO_4$ and evaporated to give an amorphous yellow foam. This foam was redissolved in MeOH and the solution adjusted to approximately pH 8 (estimated with pH paper) with $NaHCO_3$ and stirred at ambient for a further two hours. At this point the solution was evaporated and crude product purified by flash column (3 % MeOH/$CHCl_3$) to yield the title compound as a yellow foam after the removal of column solvents.

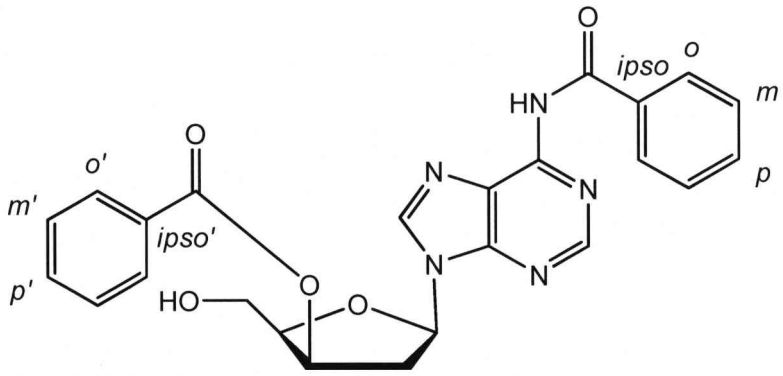
Yield: 2.2 g, 4.8 mmol, 40 %

δ^1H (400 MHz, DMSO-d_6) 2.45 (1H, m, H2'), 2.89 (1H, m, H2'), 4.38 (1H, dt, $^3J_{3'}$ = 7.50 Hz, $^3J_{5'}$ = 3.70 Hz, H4'), 4.51 – 4.57 (2H, m, H5', H3'), 4.67 (1H, dd, $^2J_{5'}$ = 11.76 Hz, $^3J_{4'}$ = 3.78 Hz, H5'), 5.85 (1H, d, $^3J_{3'}$ = 4.36 Hz, 3'-OH), 6.47 (1H, dd, ^3J = 8.34, 1.90 Hz, H1'), 7.50 – 7.57 (4H, m, *m*-, *m'*-ArH), 7.64 (2H, m, *p*-, *p'*-ArH), 7.97, 8.05 (4H, m, *o*-, *o'*- Ar-H), 8.72 (1H, s, H8), 8.76 (1H, s, H2), 11.16 (bs, NH).

δ^{13}C (100 MHz, DMSO-d_6) 41.0 (C2'), 64.4 (C5'), 69.7 (C3'), 82.5 (C4'), 83.4 (C1'), 125.9 (C5), 128.8 – 129.6, 132.8 – 133.1 (10 x Ar-C), 130.0 (*ipso'*-Ar-C), 133.7 (*ipso*-Ar-C), 143.5 (C8), 150.7 (C6), 151.9 (C2), 152.2 (C4), 166.0 (NHBz C=O, 5'-OBz C=O).

Analysis found C 62.66 H 4.61 N 15.30; $C_{24}H_{21}N_5O_5$ requires C 62.74 H 4.61 N 15.24

Found m/z (ES^+) 460.1638 ($[M + H]^+$ 100%); $[C_{24}H_{22}N_5O_5]^+$ requires 460.1621



Isolated yield from method A: 850 mg, 1.8 mmol, 15 %

δ^1H (400 MHz, DMSO-d_6) 2.98 (1H, m, H2''), 3.09 (1H, m, H2'), 3.79 (2H, m, H5'), 4.37 (1H, m, H4'), 4.94 (1H, t, $^3J_{5'}$ = 5.50 Hz, 5'-OH), 5.74 (1H, t, J = 4.08 Hz, H3'), 6.54 (1H, dd, ^3J = 7.50, 2.74 Hz, H1'), 7.50 – 7.57 (4H, m, *m*-, *m'*-Ar-H), 7.65 (2H, m, *p*-, *p'*-Ar-H), 7.81, 8.04 (4H, m, *o*-, *o'*- Ar-H), 8.60 (1H, s, H2), 8.65 (1H, s, H8), 11.13 (bs, NH).

δ^{13}C (100 MHz, DMSO-d_6) 38.1 (C2'), 59.6 (C5'), 73.4 (C3'), 83.8 (C1', C4'), 126.1 (C5), 128.8 – 129.6, 132.7 – 133.8 (10 x Ar-C), 130.0 (*ipso'*-Ar-C), 133.1 (*ipso*-Ar-C), 142.6 (C8), 150.6 (C6), 151.9 (C2), 152.2 (C4), 165.3 (*N*-benzoyl CO, 5'-O-benzoyl CO).

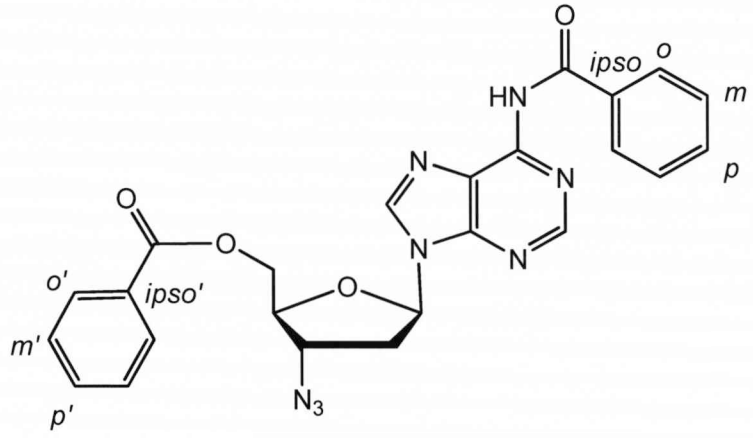
Found m/z (ES$^+$) 460.1638 ([M + H]$^+$ 100%); [$C_{24}H_{22}N_5O_5$ +Na]$^+$ requires 460.1621

Isomerisation of *N*-6, 3'-*O*-dibenzoyl-3'-*xylo*-2'-deoxyadenosine (22) to *N*-6, 5'-*O*-dibenzoyl-3'-*xylo*-2'-deoxyadenosine (18)

N-6-3'-Dibenzoyl-3'-*xylo*-2'-deoxyadenosine (**22**, 700 mg, 1.5 mmol) and $NaHCO_3$ were dissolved in dry methanol (50 ml) and stirred at ambient temperature for 6 hours. The solvent was removed under reduced pressure and the crude product purified by flash column (0 – 3 % MeOH/$CHCl_3$) to give the title compound as a clear oil.

Yield: 330 mg, 0.7 mmol, 48 %

Preparation of *N*-6, 5'-*O*-dibenzoyl-3'-azido-2'-3'-dideoxyadenosine (19)



N-6-5'-Dibenzoyl-3'-*xylo*-2'-deoxyadenosine (**18**, 3.2 g, 7.0 mmol) was dissolved in a mixture of anhydrous DCM (64 ml) and anhydrous pyridine (6.5 ml) and the solution cooled to -30 °C in a dry ice/EtOAc bath before $(TfO)_2O$ (2.9 g, 10.4 mmol) was introduced dropwise. After the addition the solution was allowed to warm to room temperature over approximately 30 minutes at which point LiN_3 (3.4 g, 70 mmol) in DMF solution (40 ml) was added as a single aliquot, turning the solution a deep red colour. The reaction was allowed to stir at ambient for 90 minutes then partitioned between water (100 ml) and $CHCl_3$ (300 ml) and the organic layer separated, washed with water (2 x 200 ml) and dried over $MgSO_4$. Following evaporation of the $CHCl_3$ the oily yellow residue was co-evaporated with toluene and with MeCN to remove traces of pyridine and DMF respectively. The crude product was purified by flash column (1 % MeOH/$CHCl_3$) and the solvents removed to give the title compound as a yellow foam.

Yield: 2.9 g, 6.0 mmol, 86 %.

δ^1H (400 MHz, DMSO-d_6) 2.72 (1H, m, H2'), 3.21 (1H, m, H2'), 4.30 (1H, pseudo-q, J = 5.05 Hz, H4'), 4.53 (1H, dd, $^2J_{5'}$ = 12.06 Hz, $^3J_{4'}$ = 5.22 Hz, H5'), 4.59 (1H, dd, $^2J_{5'}$ = 12.02 Hz, $^3J_{4'}$ = 4.26 Hz, H5'), 4.94 (1H, pseudo-q, J = 6.32 Hz, H3'), 6.51 (1H, dd, ^3J = 7.12 Hz, 5.40 Hz, H1'), 7.54 (4H, m, *m*-, *m'*-ArH), 7.66 (2H, m, *p*-, *p'*-Ar-H), 7.91, 8.04 (4H, 2 x m, *o*-, *o'*--Ar-H), 8.71 (1H, s, H8), 8.65 (1H, s, H2), 11.22 (bs, NH).

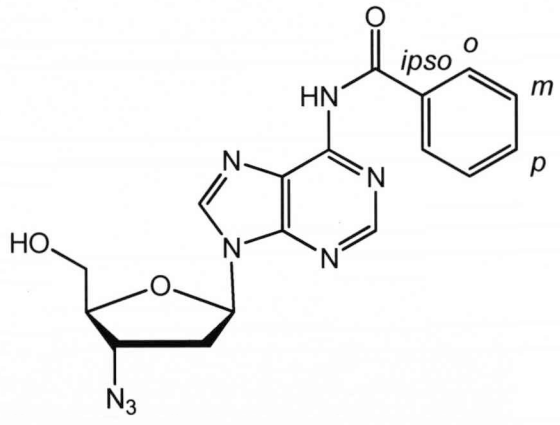
$\delta^{13}C$ (100 MHz, DMSO-d_6) 36.0 (C2'), 61.7 (C3'), 64.1 (C5'), 81.7 (C4'), 83.6 (C1'), 126.3 (C5), 128.8, 128.9, 129.1, 129.6 (*o, o', m, m'*-Ar-C), 129.56 (*ipso'*-Ar-C), 132.8, 133.7 (*p, p'*-Ar-C), 133.9 (*ipso*-Ar-C), 143.7 (C8), 150.8 (C6), 152.0 (C2), 152.1 (C4), 165.8 (5'-OBz C=O) 166.0 (NHBz C=O).

Analysis found C 53.65 H 4.21 N 29.53; $C_{24}H_{21}N_8O_4$ requires C 53.68 H 4.24 N 29.46

Found m/z (ES^+) 485.1686 ($[M + H]^+$ 100%); $[C_{24}H_{21}N_8O_4 + H]^+$ requires 485.1686

IR ν = 2102.0 cm^{-1} (azide asymmetric stretch)

Preparation of *N*-6-benzoyl-3'-azido-2'-3'-dideoxyadenosine (20)



N-6-5'-Dibenzoyl-3'-azido-2'-3'-dideoxyadenosine (**19**, 5.7 g, 11.8 mmol) was dissolved in a mixture of 1,4-dioxane/MeOH/water (210 ml, 5:4:1) and the solution cooled to 0 °C. 1 M NaOH (25 ml) was added and the reaction stirred for a further 30 minutes before being neutralised with Amberlite® H^+ ion exchange resin. The resin was removed by filtration and the reaction solvents evaporated under reduced pressure to give crude product as a yellow oil. This yellow oil was purified by flash column (0 – 4 % MeOH/$CHCl_3$) to give the desired product as a fine white powder after removal of the solvents. Recrystallisation from MeOH changed the physical form to a crumbly white powder.

Yield: 2.2 g, 5.8 mmol, 49 %, m.pt. 181-183 °C

δ^1H (400 MHz, DMSO-d_6) 2.20 (1H, m, H2'), 3.05 (1H, m, H2'), 3.61 (2H, m, H5'), 3.97 (1H, pseudo-q, J = 6.92 Hz, H4'), 4.68 (1H, dt, $^3J_{4'}$ = 6.64 Hz, $^3J_{2'}$ = 5.26 Hz, H3'), 5.21 (1H, t, $^3J_{5'}$ = 5.6 Hz 5'-OH), 6.47 (1H, t, $^3J_{2'}$ = 6.26 Hz, H1'), 7.56 (2H, m, *m*-Ar-H), 7.64 (1H, m, *p*-Ar-H), 8.05 (2H, m, *o*-Ar-H), 8.71 (1H, s, H8), 8.65 (1H, s, H2), 11.24 (bs, NH).

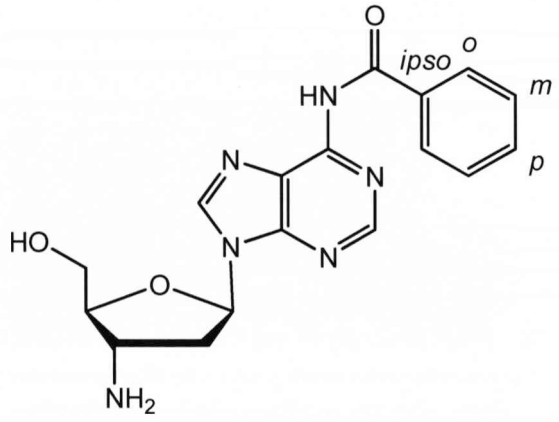
δ^{13}C (100 MHz, DMSO-d_6) 36.5 (C2'), 61.0 (C3'), 61.4 (C5'), 83.6 (C4'), 85.1 (C1'), 126.2 (C5), 128.9 (*o*, *m*-Ar-C), 132.8 (*p*-Ar-C), 133.7 (*ipso*-Ar-C), 143.4 (C8), 150.7 (C6), 152.0 (C2), 152.2 (C4), 166.0 (C=O).

Analysis found C 53.65 H 4.21 N 29.53; $C_{17}H_{16}N_8O_3$ requires C 53.68 H 4.24 N 29.46

Found m/z (ES$^+$) 381.1430 ([M + H]$^+$ 100%); [$C_{17}H_{17}N_8O_3$ + H]$^+$ requires 381.1424

IR ν = 2096.2 cm^{-1} (azide asymmetric stretch)

Preparation of *N*-6-benzoyl-3'-amino-2'-3'-dideoxyadenosine (24)



N-6-Benzoyl-3'-amino-2'-3'-dideoxyadenosine (**20**, 400 mg, 1.0 mmol) and Pd/C was suspended in MeOH/H_2O (2:1, 8 ml) and stirred vigorously under an H_2 atmosphere for 2 hours at ambient temperature. Pd/C was removed by filtration through Celite and the filter medium washed with 30 % aqueous MeOH and the solvents removed under reduced pressure to give the desired compound as an amorphous white foam.

Yield: 372 mg, 1.0 mmol, > 99 %

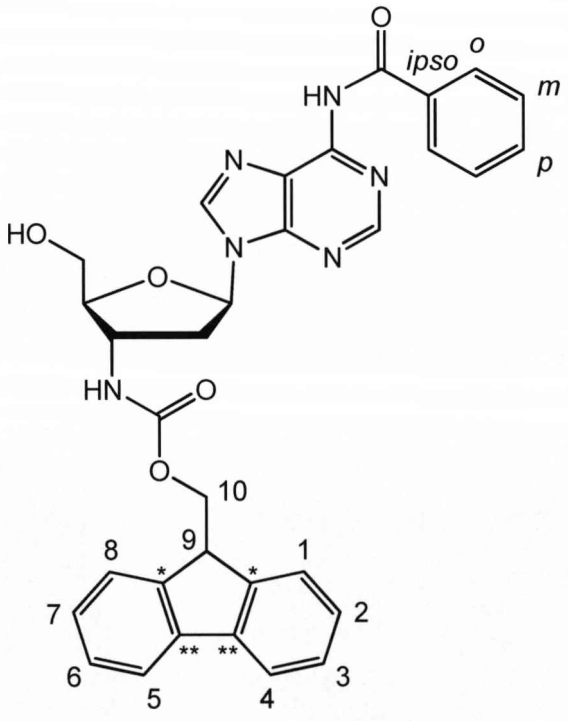
δ^1H (400 MHz, DMSO-d_6) 2.28 (1H, m, H2'), 2.66 (1H, m, H2'), 3.56 (1H, m, H4'), 3.67 (3H, m, H3', H5'), 4.98 (1H, bs, 5'-OH), 6.45 (1H, dd, $^3J_{2'}$ = 6.98 Hz, $^3J_{2'}$ = 4.42 Hz, H1'), 7.54 (2H, t, ^3J = 7.48 Hz, *m*-Ar-H), 7.63 (1H, t, ^3J_m = 7.36 Hz, *p*-Ar-H), 8.05 (2H, m, *o*-Ar-H), 8.67 (1H, s, H8), 8.72 (1H, s, H2), 11.24 (bs, NH).

δ^{13}C (100 MHz, DMSO-d_6) 41.0 (C2'), 51.8 (C3'), 61.7 (C5'), 83.7 (C1'), 88.6 (C4'), 126.2 (C5), 128.7, 128.8 (*o*, 2-Ar-C), 132.6 (*p*-Ar-C), 134.2 (*ipso*-Ar-C), 143.0 (C8), 151.2 (C6), 151.8 (C2), 152.0 (C4), 166.2 (C=O).

Analysis found C 57.30 H 5.07 N 23.55; $C_{17}H_{18}N_6O_3$ requires C 57.62 H 5.12 N 23.72

Found m/z (ES$^+$) 377.1329 ([M + Na]$^+$ 100%); [$C_{17}H_{18}N_6O_3$ + Na]$^+$ requires 377.1338

Preparation of *N*-6-benzoyl-3'-*N*-(9-fluoromethoxycarbonyl)-2'-3'-dideoxyadenosine (25)



N-6-Benzoyl-3'-amino-2'-3'-dideoxyadenosine (**24**, 200 mg, 0.56 mmol) was dissolved in 1,4-dioxane : water (1 : 1, 4 ml) and K_2CO_3 (156 mg, 1.12 mmol) added. After stirring for 30 minutes at ambient temperature the solution was cooled to 0 °C in an ice bath. FmocCl (161 mg, 0.62 mmol) was added and the reaction stirred at 0 °C for 6 hours. TLC analysis showed an incomplete reaction after this time, therefore, further FmocCl (30 mg, 0.11 mmol) was added to the now cloudy white suspension and stirred for 30 minutes at 0 °C. Water (10 ml) was added, causing more white solid product to precipitate out which was collected by filtration, washed thoroughly with water and purified by flash column (2 – 4 % MeOH/$CHCl_3$) to give product as a white foam after removal of chromatography solvents.

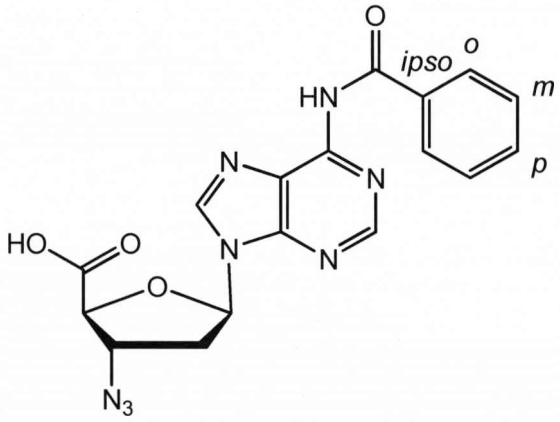
Yield: 256 mg, 0.44 mmol, 79 %

δ^1H (400 MHz, DMSO-d_6) 2.47 (1H, m, H2'), 2.84 (1H, m, H2'), 3.55, 3.67 (2H, m, H5'), 3.95 (1H, pseudo-t, H4'), 4.26 (1H, t, $^3J_{10}$ = 6.68 Hz Fmoc H-9), 4.38 (3H, m, H3', Fmoc H-10), 5.05 (1H, t, $^3J_{5'}$ = 5.46 Hz, 5'-OH), 6.51 (1H, t, $^3J_{2'}$ = 6.26 Hz, H1'), 7.36 (2H, t, ^3J = 7.44 Hz, Fmoc H-2, 7), 7.43 (2H, t, ^3J = 7.50 Hz, Fmoc H-3, 6), 7.56 (2H, t, ^3J = 7.54 Hz, *m*-Ar-H), 7.63 (1H, m, *p*-Ar-H), 7.72 (1H, d, $^3J_{2,7}$ = 7.48 Hz, Fmoc H-1, 8), 7.83 (1H, d, $^3J_{3'}$ = 7.60 Hz, 3'-NH), 7.90 (2H, d, $^3J_{3,6}$ = 7.32 Hz, Fmoc H-4, 5), 8.05 (2H, d, ^3J_m = 7.40 Hz, *o*-Ar-H), 8.72 (1H, s, H8), 8.77 (1H, s, H2).

δ^{13}C (100 MHz, DMSO-d_6) 37.6 (C2'), 47.1 (Fmoc C-9), 51.4 (Fmoc C-10), 61.9 (C5'), 65.8 (C3'), 83.7 (C1'), 85.9 (C4'), 120.5 (Fmoc C-4, 5), 125.5 (Fmoc C-1, 8), 126.2 (C5), 127.4 (Fmoc C-2, 7), 128.0 (Fmoc C-3, 6), 128.8 (*o*, *m*-Ar-C), 132.8 (*p*-Ar-C), 133.6 (*ipso*-Ar-C), 141.1 (Fmoc-C**), 143.2 (C8), 144.2 (Fmoc-C*), 150.7 (C6), 151.9 (C2), 152.2 (C4), 156.2 (Fmoc C=O), 166.0 (NHBz C=O).

Found m/z (ES$^+$) 599.2035 ([M + Na]$^+$ 100%); [$C_{32}H_{28}N_6O_5$ + Na]$^+$ requires 599.2019

Preparation of *N*-6-benzoyl-3'-azido-2'-3'-dideoxyadenosine-5'-carboxylic acid (26)



N-6-Benzoyl-3'-azido-2'-3'-dideoxyadenosine (**20**, 800 mg, 2.1 mmol) was suspended in MeCN/H_2O (1 : 1, 13 ml) and TEMPO (329 mg, 2.1 mmol) and BAIB (1.5 g, 4.6 mmol) were added. After 2 hours a white precipitate had appeared and no starting material remained by TLC, therefore, Et_2O (25 ml) was added to the suspension and the powdery precipitate isolated by filtration.

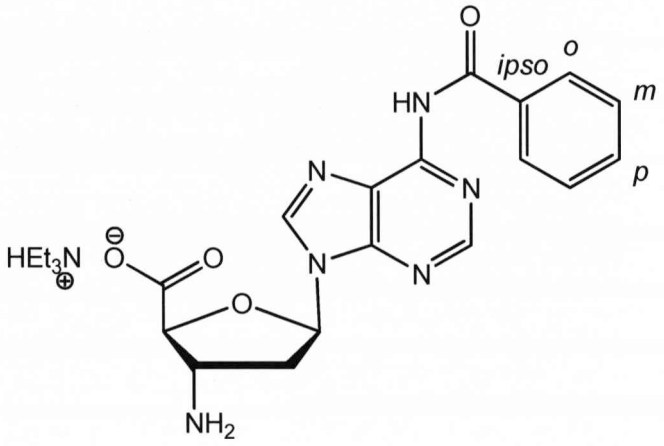
Yield: 682 mg, 1.7 mmol, 82 %

δ^1H (400 MHz, DMSO-d_6) 2.65 (1H, m, H2'), 3.03 (1H, m, H2'), 4.60 (1H, d, $^3J_{3'}$ = 4.08 Hz, H4'), 5.06 (1H, dt, $^3J_{4'}$ = 4.65 Hz, $^3J_{2'}$ = 6.50 Hz, H3'), 6.58 (1H, m, H1'), 7.56 (2H, m, *m*-Ar-H), 7.65 (1H, m, *p*-Ar-H), 8.05 (2H, m, *o*-Ar-H), 8.70 (1H, s, H8), 8.76 (1H, s, H2), 11.23 (bs, NH).

δ^{13}C (100 MHz, DMSO-d_6) 36.4 (C2'), 63.5 (C3'), 82.1 (C4'), 84.2 (C1'), 126.0 (C5), 128.8 (*o*, *m*-Ar-C),132.8 (*p*-Ar-C), 133.7 (*ipso*-Ar-C), 143.3 (C8), 150.7 (C6), 152.1 (C2), 152.4 (C4), 166.0 (NHBz C=O), 171.4 (C5').

Found m/z (ES^-) 393.1042 ($[M - H]^-$ 100%); $[C_{17}H_{14}N_8O_4 - H]^-$ requires 393.1060

Preparation of *N*-6-benzoyl-3'-amino-2'-3'-dideoxyadenosine-5'-carboxylic acid triethylamine salt (27)



N-6-Benzoyl-3'-azido-2'-3'-dideoxyadenosine-5'-carboxylic acid (**26**, 605 mg, 1.5 mmol) was dissolved in anhydrous 15 % Et_3N/pyridine (6 ml), cooled to 0 °C and H_2S bubbled through the solution for 5 minutes. The now brown solution was stirred for a further 30 minutes before H_2O (5 ml) was poured in causing a precipitate to appear. This precipitate was removed by centrifugation and the remaining supernatant evaporated under reduced pressure to give the desired product as a pale yellow amorphous solid.

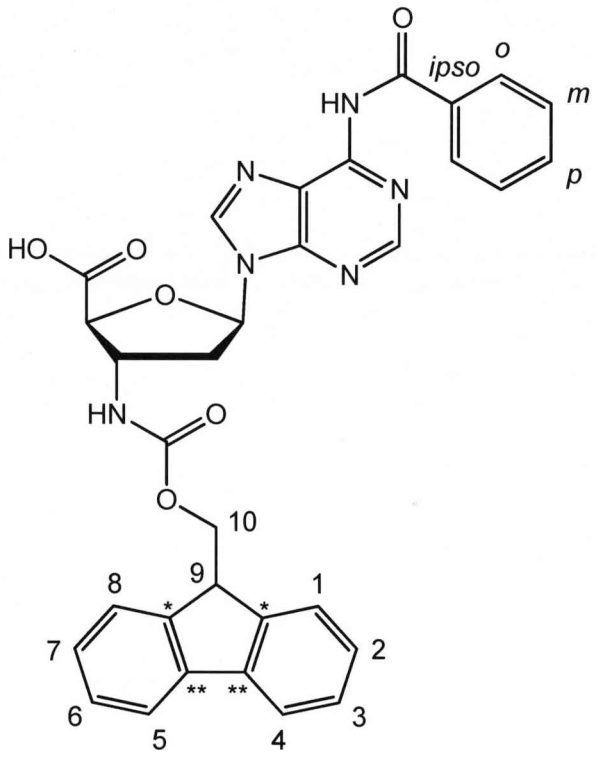
Yield: 405 mg, 0.86 mmol, 56 %

δ^1H (400 MHz, DMSO-d_6) 1.09 (3H, t, ^3J = 7.24 Hz, Et_3N CH_3), 2.55 (1H, m, H2'), 2.70 (1H, m, H2'), 2.87 (2H, q, ^3J = 7.25 Hz, Et_3N CH_2), 4.04 (1H, m, H3'), 4.31 (1H, d, $^3J_{3'}$ = 3.32 Hz, H4'), 6.69 (1H, t, $^3J_{2'}$ = 6.54 Hz, H1'), 7.55 (2H, t, ^3J = 7.84 Hz, *m*-Ar-H), 7.63 (1H, m, *p*-Ar-H), 8.04 (2H, m, *o*-Ar-H), 8.73 (1H, s, H2), 9.23 (1H, s, H8).

δ^{13}C (100 MHz, DMSO-d_6) 9.6 (Et_3N CH_3) 37.5 (C2'), 45.6 (Et_3N CH_2) 54.9 (C3'), 84.0 (C1'), 84.4 (C4'), 126.0 (C5), 128.1 (*o*, *m*-Ar-C),132.8 (*p*-Ar-C), 133.8 (*ipso*-Ar-C), 144.2 (C8), 150.4 (C6), 151.8 (C2), 152.3 (C4), 166.0 (NHBz C=O), 172.5 (C5').

Found m/z (ES^+) 369.1301 ($[M + H]^+$ 52 %); $[C_{17}H_{16}N_6O_4 + H]^+$ requires 369.1311

Preparation of *N*-6-benzoyl-3'-*N*-fmoc-amino-2'-3'-dideoxyadenosine-5'-carboxylic acid (23)



N-6-Benzoyl-3'-amino-2'-3'-dideoxyadenosine-5'-carboxylic acid triethylamine salt (**27**, 395 mg, 0.84 mmol) was dissolved in 1,4-dioxane/H_2O (2:1, 8 ml) containing K_2CO_3 (116 mg, 0.84 mmol). The solution was cooled to 0 °C before FmocCl (260 mg, 1.0 mmol was added in a single portion and the solution stirred at 0 °C for 1 hour. The solution was then acidified to pH 2 with $KHSO_4$ causing the product to precipitate out of solution as a white powdery solid which was isolated by filtration.

Yield: 292 mg, 0.5 mmol, 59 %

δ^1H (400 MHz, DMSO-d_6) 2.56 (1H, m, H2'), 2.92 (1H, m, H2'), 4.27 (1H, m, Fmoc H-9), 4.38 (2H, m, Fmoc H-10), 4.43 (1H, m, H4'), 4.64 (1H, m, H3'), 6.67 (1H, m, H1'), 7.35 (2H, m, Fmoc H-2, 7), 7.42 (2H, m, Fmoc H-3, 6), 7.54 (2H, m, *m*-Ar-H), 7.64 (1H, m, *p*-Ar-H), 7.74 (1H, m, Fmoc H-1, 8), 7.84 (1H, d, $^3J_{3'}$ = 7.40 Hz, NH), 7.90 (2H, m, Fmoc H-4, 5), 8.05 (2H, m, *o*-Ar-H), 8.75 (1H, s, H8), 8.77 (1H, s, H2).

$\delta^{13}C$ (100 MHz, DMSO-d_6) 36.1 (C2'), 46.6 (Fmoc C-9), 54.6 (C3'), 65.5 (Fmoc C10), 82.1 (C1'), 84.0 (C4'), 120.0 (Fmoc C-4, 5), 125.0 (Fmoc C-1, 8), 127.0 (C5), 127.5 (Fmoc C-2, 7), 128.3 (*o*, *m*-Ar-C), 128.8 (Fmoc C-3,6), 132.3 (*p*-Ar-C), 133.4 (*ipso*-Ar-C), 140.6 (Fmoc-C**), 143.8 (C8), 144.3 (Fmoc-C*), 149.9 (C6), 151.2 (C2), 151.8 (C4), 155.5 (Fmoc C=O), 172.2 (C5').

Found m/z (ES^+) 589.1833 ($[M - H]^-$ 100%); $[C_{32}H_{25}N_6O_6 - H]^-$ requires 589.1836

4.3 Manual Solid Phase Peptide Synthesis.

Peptides were synthesised manually in a batch procedure. *N*-Fmoc-9-amino-xanthe-3-yloxymethy-Sieber amide resin (0.71 mmol/g NHFmoc) was purchased from Novabiochem.

NMR Spectroscopy

^{1}H NMR spectra were recorded on either a Varian Unity Inova 500 MHz or a Bruker DRX 500 MHz machine by Julie Fisher and co-workers, University of Leeds.

TNBS Test.

The TNBS test was performed using the following reagents:

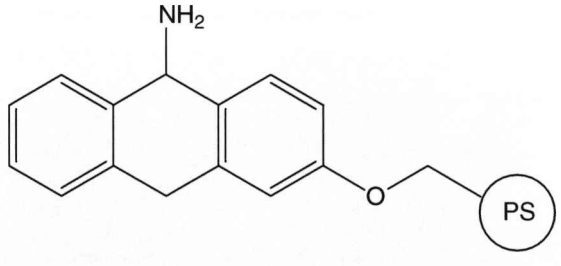
Reagent A: 10 % v/v DIPEA in DMF

Reagent B: 1 % v/v TNBS in DMF

Approximately 10 resin beads were extracted from the drained resin bed and placed in a microcentrifuge tube. One drop of each reagent solution was added to the beads and the tube left to stand at ambient temperature for five minutes, then the colour of the beads noted; red - positive (free amino groups present; successful deprotection or failed coupling steps), yellow – negative (no free amino groups present; unsuccessful deprotection or successful coupling steps).

General procedure for solid phase synthesis of 5'-amide peptide oligomers:

Couplings were performed on Sieber amide resin (200 mg, 142 μmol NHFmoc) loaded at 25 % capacity using the following quantities:



Coupling	Component	Quantity
1	Fmoc amino acid	35.5 µmol
	0.2 M HBTU	178 µl
	DIPEA	6.8 µl, 39 µmol
2 onwards	Fmoc amino acid	50 µmol
	0.2 M HBTU	250 µl
	DIPEA	9.5 µl, 55 µmol

Swelling and deprotection of Sieber amide resin.

The bed of Sieber amide resin was swelled overnight with three times resin bed volume of DCM. The swelling solvent was removed and the resin washed with DMAc/DMF (x 5) and drained. Deprotection solution (2 % v/v DBU in DMAc/DMF) was applied to the resin to remove the Fmoc protecting group. The resin bed was covered with the deprotection solution and the reaction vessel agitated on an orbital shaker for 5 minutes. The deprotection solution was removed and fresh solution re-applied a further two times. A TNBS test was performed to confirm the removal of the protecting group. The resin bed was then washed thoroughly with DMAc/DMF (x 5).

Coupling of the first nucleoside monomer residue.

The nucleoside β-amino acid (35.5 µmol) was dissolved in DMAc/DMF (178 µl) and pre-activated for 60 s with HBTU in DMAc/DMF (0.2 M, 178 µl) and DIPEA (6.8 µl, 39.0 µmol). The deprotected resin bed was then treated with this solution and agitated at ambient temperature for five hours. Following removal of the coupling solution the resin bed was washed with DMAc/DMF (x 5) and capped as below.

Capping unreacted sites.

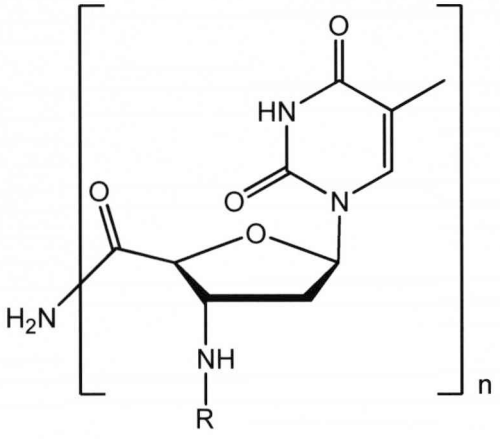
Capping solution (Ac_2O, DMAc/DMF, collidine, 1:8:1) was added to cover the resin bed and the reaction vessel agitated for 5 minutes on an orbital shaker. The resin bed was then washed thoroughly with DMAc/DMF (x 5).

Further residue coupling.

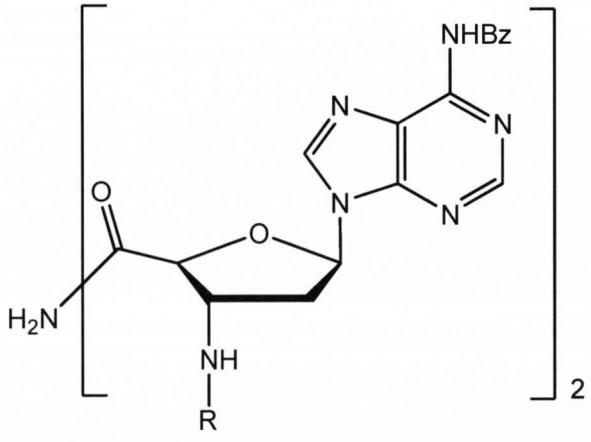
The cycle of deprotection, coupling and capping was continued according to the above procedure for further nucleoside residues. For residues after the first,

coupling times were reduced to 30 min and a TNBS test performed to confirm successful couplings. Quantities for subsequent couplings are shown in the tables above. When all the required amino acid couplings had been performed, to obtain Fmoc off compounds the resin was subjected to a further deprotection step. To obtain Fmoc on compounds this extra deprotection was omitted from the procedure. The resin bed was dried by washing with DMAc/DMF (x 5), DCM (x 5), MeOH (x 5) and vacuum applied to the reaction vessel for 30 minutes. Finally the vessel was stored in a vacuum dessicator for 24 hours.

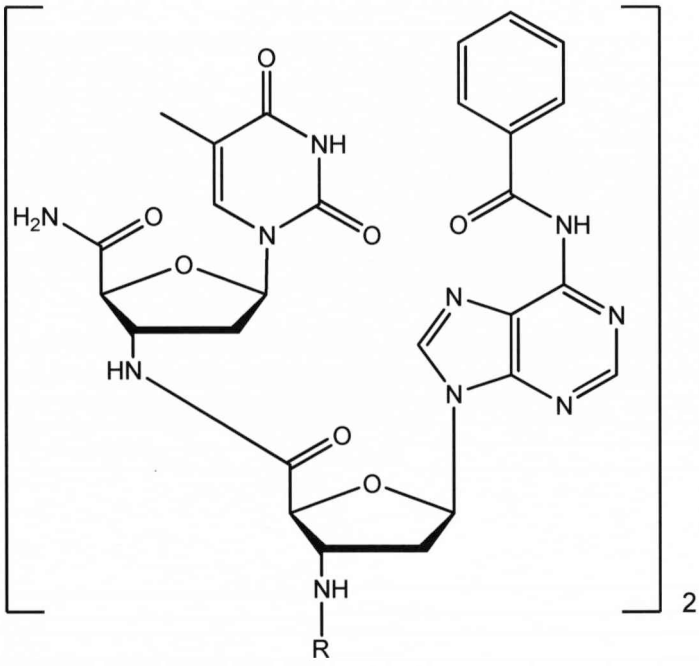
Cleavage of oligomer from solid support.



Oligomer	n	R
10	2	H
11	4	H
12	4	Fmoc
13	8	H
14	8	Fmoc



Oligomer	n	R
28	4	Fmoc
29	4	H



Oligomer	Sequence	R
30	TATA	Fmoc
31	TATA	H

The dry resin was swelled in DCM for 30 minutes then washed three times with fresh DCM. The resin was covered with cleavage solution (1 % v/v TFA in DCM) and agitated for 2 minutes on an orbital shaker. The cleavage mixture was removed and the procedure repeated nine further times. Residual peptide was washed from the resin bed with DCM (x 3) and MeOH (x 3) then the combined cleavage solutions evaporated to dryness under reduced pressure. The peptide was then triturated with Et_2O and centrifuged to form a pellet which was dried under a flow of nitrogen gas.

HPLC Retention Times

Thymidine peptides

Oligomer	n	R	Retention Time/mins
10	2	H	17.88
11	4	H	21.59
12	4	Fmoc	37.49
13	8	H	25.18
14	8	Fmoc	34.88

2'-Deoxyadenosine peptides

Oligomer	n	R	Retention Time/mins
28	4	Fmoc	N/A
29	4	H	27.66

Mixed sequence peptides

Oligomer	Sequence	R	Retention Time/mins
30	TATA	Fmoc	N/A
31	TATA	H	21.40

Selected NMR data

Peptide 12

	NH	NH2	H1'	H2'	H3'	H4'	H-6	OMe
1	8.85	7.60	6.48	2.41/2.33	4.60	4.35	8.06	1.90
2	8.92	N/A	6.33	2.46/2.39	4.64	4.38	8.06	1.90
3	8.94	N/A	6.33	2.46/2.34	4.66	4.39	8.21	1.90
4	8.13	N/A	6.32	2.43/2.30	4.44	4.40	8.06	1.90

4.4 Thermal Melting Studies

UV melting curves were recorded on a Hewlett-Packard 8452A diode array spectrophotometer equipped with a peltier device. dA_8 and T_8 DNA strands were synthesised by James Gaynor. Solutions of the following were made up in 10 mM NaH_2PO_4 buffer containing 100 mM NaCl adjusted to pH 7.0 with HCl:

Solution	Components
A	dT_8 (6 μM), dA_8 (6 μM)
B	Peptide (**13**) (3 μM), dA_8 (3 μM)
C	Peptide (**12**) (6 μM)
D	Peptide (**13**) (3 μM)

Solutions were heated to 90 °C for 3 minutes then cooled gradually.

4.5 Fluorous-Phase Synthesis

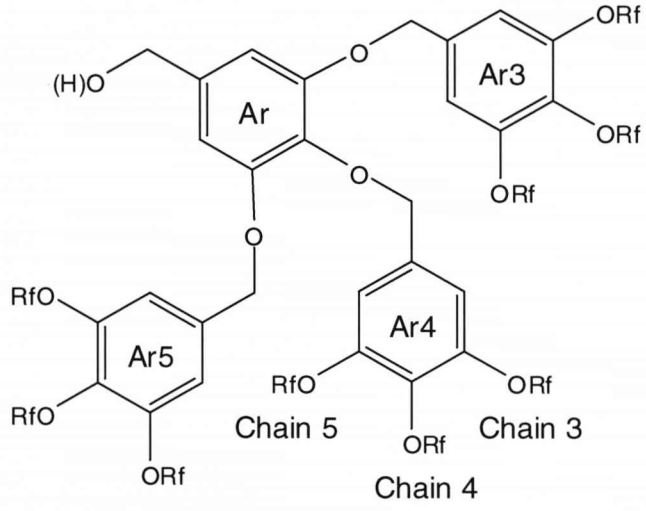
Reagents

FC-72 was purchased from 3M (Japan) and Acota (UK) Ltd.

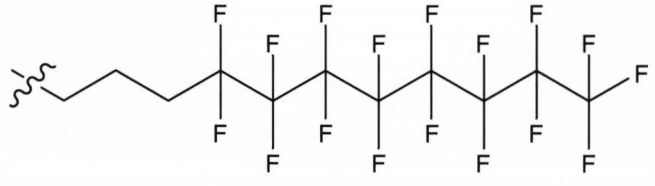
THF: Distilled from sodium/benzophenone.

MNTP was prepared by the group of Associate Prof. Takeshi Wada, Graduate School of Frontier Sciences, University of Tokyo.[275]

FD-O(H) refers to the fluorous dendrimer shown below and was also prepared by the group of Associate Prof. Takeshi Wada:



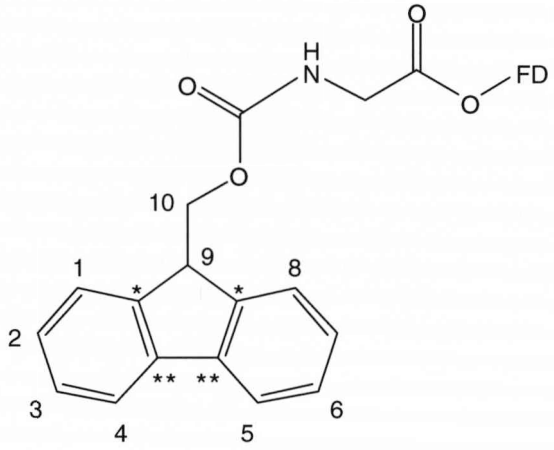
Rf =



Mass Spectrometry

MALDI-TOF Mass Spectra were obtained by the EPSRC National Mass Spec Centre, Swansea on a Voyager DE-STR MALDI-TOF spectrometer using a pentafluorobenzoic acid matrix with NaOAc additive.

Preparation of FD-O-*N*-Fmoc-Glycine (33)



Fmoc-glycine (30 mg, 0.1 mmol) was dissolved in anhydrous THF (1 ml) and FD-O(H) (470 mg, 0.1 mmol) was suspended in this solution. To this mixture was added 2,4,6-collidine (65 µl, 0.5 mmol) and MNTP (108 mg, 0.24 mmol) and the reaction heated to 45 °C and stirred at this temperature for 1 hour. After cooling to ambient, the now biphasic solution was washed with water (5 ml) and $CHCl_3$ (5 ml) and both the aqueous and organic phases backwashed with FC-72 (10 x 5ml). The combined fluorous phases were then evaporated to give the desired product as an amorphous white solid.

Yield: 440 mg, 0.88 mmol, 88 %

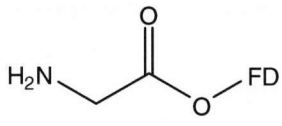
δ^1H (400 MHz, $CDCl_3/C_6F_6$, (3:1)) 2.04 – 2.11 (18H, m, $CF_2CH_2\underline{CH_2}$), 2.31 – 2.38 (18H, m, $CF_2\underline{CH_2}$), 3.82 (4H, t, ^3J = 5.60 Hz, OCH_2 (Ar3, Ar5, Chain 4), 3.92 – 4.04 (14H, m, OCH_2, Ar3 (Chains 3, 5), Ar4, (Chains 3, 4, 5) Ar5, (Chains 3, 5), 4.13 (2H, d, $^3J_{H\text{-}9}$ = 5.68 Hz, Fmoc H-10), 4.43 (2H, d, ^3J = 7.2 Hz, Gly-CH_2) 4.96 (2H, s, Ar-O-$\underline{CH_2}$-Ar4), 5.06 (4H, s, Ar-O-$\underline{CH_2}$-Ar3, Ar-O-$\underline{CH_2}$-Ar5), 5.24 (2H, s, Ar-$\underline{CH_2}$-O-Gly), 5.27 (1H, t, $^3J_{H\text{-}10}$= 5.78 Hz, Fmoc H-9), 6.68 (2H, s, *o*-Ar4), 6.71 (4H, s, *o*-Ar3, *o*-Ar5), 6.80 (2H, s, *o*-Ar), 7.30 (2H, dt, J= 7.68 Hz, J= 0.84 Hz, Fmoc H-2, 7), 7.39 (2H, m, Fmoc H-3, 6), 7.60 (1H, m, Fmoc H-1, 8), 7.58 (2H, m, Fmoc H-4, 5).

δ^{13}C (100 MHz, $CDCl_3/C_6F_6$, (3:1)) 19.9 ($CF_2CH_2\underline{CH_2}$), 27.0 ($CF_2\underline{CH_2}$), 67.0 (Gly-O-$\underline{C}H_2$-Ar), Ar3-O-$\underline{CH_2}$-CH_2, Ar5-O-$\underline{CH_2}$-CH_2 (chain 4)), 67.6 – 72.1 (Ar3, Ar5-O-$\underline{CH_2}$-CH_2 (chains 3 and 5), Ar4-O-$\underline{CH_2}$-CH_2 (chains3, 4, 5), 68.1 (Gly- $\underline{C}H_2$), 105.7 –

106.1 (*o*-ArC, *o*-Ar3-C, *o*-Ar4-C, *o*-Ar5-C), 120.0 (Fmoc C-4, 5), 125.2 (Fmoc C-1, 8), 127.3 (Fmoc C-2, 7), 128.0 (Fmoc C-3, 6), 138.0 (*p*-ArC, *p*-Ar3-C, *p*-Ar4-C, *p*-Ar5-C), 140.9 (Fmoc-C**), 143.5 (Fmoc-C*), 152.0 (*m*-ArC, *m*-Ar3-C, *m*-Ar4-C, *m*-Ar5-C)

Found (MALDI-TOF) 5013.3 ($[M + Na]^+$ 100%); 5029.3 ($[M + K]^+$ 34%);

Preparation of FD-O-Glycine (34)



FD-O-N-Fmoc-glycine (**33**, 110 mg, 22 µmol) was dissolved in FC-72 (1.0 ml) and 10 % piperidine in DMF solution (1.0 ml) added. The biphasic mixture was stirred at ambient temperature for 30 minutes before MeCN (5 ml) was added and the layers separated. The organic layer was extracted with further FC-72 (5 x 5 ml) and all the fluorous phases combined and evaporated to give product as an off-white amorphous foam.

Yield 70 mg, 14.5 µmol, 66 %

δ^1H (400 MHz, $CDCl_3/C_6F_6$, (3:1)) 2.04 – 2.12 (18H, m, $CF_2CH_2\underline{CH_2}$), 2.31 – 2.37 (18H, m, $CF_2\underline{CH_2}$), 3.56 (2H, s, Gly-CH_2), 3.81 (4H, t, ^3J = 5.56 Hz, OCH_2 (Ar3, Ar5, Chain 4), 3.94 – 4.04 (14H, m, OCH_2, Ar3 (Chains 3, 5), Ar4, (Chains 3, 4, 5) Ar5, (Chains 3, 5), 4.98 (2H, s, Ar-O-$\underline{CH_2}$-Ar4), 5.07 (4H, s, Ar-O-$\underline{CH_2}$-Ar3, Ar-O-$\underline{CH_2}$-Ar5), 5.18 (2H, s, Ar-$\underline{CH_2}$-O-Gly), 6.68 (2H, s, *o*-Ar4), 6.72 (4H, s, *o*-Ar3, *o*-Ar5), 6.80 (2H, s, *o*-Ar).

Found (MALDI-TOF) 5013.3 ($[M + Na]^+$ 100%); 5029.3 ($[M + K]^+$ 34%);

5 Bibliography

1 Meischer, F. *Die Histochemischen und Physiologischen Arbeiten: Leipzig* **1897**.

2 Hertwig, O. *Jen. Z. Natur.* **1884**.

3 Levene, P. S.; London, E. S. *J. Biol. Chem.* **1929**, *83*, 793.

4 Franklin, R. E.; Gosling, R. G. *Trans. Faraday Soc.* **1954**, *50*, 298.

5 Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.

6 Olson, W. K. *Biopolymers* **1973**, *12*, 1787.

7 Olson, W. K.; Sussman, J. L. *J. Am. Chem. Soc.* **1982**, *104*, 270.

8 Ohyama, T. *DNA Conformation and Transcription*; Springer-Verlag: New York, **2005**.

9 Zamenhof, S.; Brawerman, G.; Chargaff, E. *Biochim. Biophys. Acta* **1952**, *9*, 402.

10 Blackburn, G. M.; Gait, M. J.; Loakes, D.; Williams, D. M. *Nucleic Acids in Chemistry and Biology*; 3rd ed.; RSC Publishing, **2006**.

11 http://en.wikipedia.org/wiki/Image:DNA_Overview.png

12 Conner, B. N.; Takano, T.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Nature* **1982**, *295*, 294.

13 Dickerson, R. E.; Drew, H. R. *J. Mol. Biol.* **1981**, *149*, 761.

14 Belmont, P.; Constant, J. F.; Demeunynck, M. *Chem. Soc. Rev.* **2001**, *30*, 70.

15 Rich, A.; Zhang, S. G. *Nat. Rev. Genet.* **2003**, *4*, 566.

16 Rich, A.; Nordheim, A.; Wang, A. H. J. *Ann. Rev. Biochem.* **1984**, *53*, 791.

17 Drew, H.; Takano, T.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Nature* **1980**, *286*, 567.

18 Moller, A.; Nordheim, A.; Kozlowski, S. A.; Patel, D. J.; Rich, A. *Biochemistry* **1984**, *23*, 54.

19 Leslie, A. G. W.; Arnott, S.; Chandrasekaran, R.; Ratliff, R. L. *J. Mol. Biol.* **1980**, *143*, 49.

20 Saenger, W. *Principles of Nucleic Acid Structure*; Springer Verlag: New York, **1984**.

21 Wahl, M. C.; Sundaralingam, M. *Biopolymers* **1997**, *44*, 45.

22 Lu, X. J.; Shakked, Z.; Olson, W. K. *J. Mol. Biol.* **2000**, *300*, 819.

23 Oh, D. B.; Kim, Y. G.; Rich, A. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 16666.

24 Hoogsteen, K. *Acta Crystallographica* **1963**, *16*, 907.

25 Sun, J. S.; Garestier, T.; Helene, C. *Curr. Op. Struc. Biol.* **1996**, *6*, 327.

[26] Holliday, R. *Genet. Res.* **1964**, *5*, 282.

[27] Lilley, D. M. J. *Q. Rev. Biophys.* **2000**, *33*, 109.

[28] Liu, Y. L.; West, S. C. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 937.

[29] Davis, J. T. *Angew. Chem. Int. Ed.* **2004**, *43*, 668.

[30] Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. *Nucleic Acids Res.* **2006**, *34*, 5402.

[31] Huppert, J. L.; Balasubramanian, S. *Nucleic Acids Res.* **2005**, *33*, 2908.

[32] Blackburn, E. H. *Nature* **1991**, *350*, 569.

[33] Zahler, A. M.; Williamson, J. R.; Cech, T. R.; Prescott, D. M. *Nature* **1991**, *350*, 718.

[34] Mergny, J. L.; Riou, J. F.; Mailliet, P.; Teulade-Fichou, M. P.; Gilson, E. *Nucleic Acids Res.* **2002**, *30*, 839.

[35] Gehring, K.; Leroy, J. L.; Gueron, M. *Nature* **1993**, *363*, 561.

[36] Mergny, J. L.; Lacroix, L.; Han, X. G.; Leroy, J. L.; Helene, C. *J. Am. Chem. Soc.* **1995**, *117*, 8887.

[37] Catasti, P.; Chen, X.; Deaven, L. L.; Moyzis, R. K.; Bradbury, E. M.; Gupta, G. *J. Mol. Biol.* **1997**, *272*, 369.

[38] Han, X. G.; Leroy, J. L.; Gueron, M. *J. Mol. Biol.* **1998**, *278*, 949.

[39] Gueron, M.; Leroy, J. L. *Curr. Op. Struc. Biol.* **2000**, *10*, 326.

[40] Astbury, W. T.; Street, A. *Trans. R. Soc. Lond.* **1931**, *A320*, 75.

[41] Pauling, L.; Corey, R. B.; Branson, H. R. *Proc. Natl. Acad. Sci. USA* **1951**, *37*, 205.

[42] Astbury, W. T. *Trans. Faraday Soc.* **1933**, *29*, 193.

[43] Pauling, L.; Corey, R. B. *Nature* **1951**, *168*, 550.

[44] Fodje, M. N.; Al-Karadaghi, S. *Protein Eng.* **2002**, *15*, 353.

[45] Huggins, M. *Chem. Rev.* **1943**, *32*, 195.

[46] Bodkin, M. J.; Goodfellow, J. M. *Protein Sci.* **1995**, *4*, 603.

[47] Scholtz, J. M.; Baldwin, R. L. *Ann. Rev. Biophys. Biomolec. Struct.* **1992**, *21*, 95.

[48] Shepherd, N. E.; Hoang, H. N.; Abbenante, G.; Fairlie, D. P. *J. Am. Chem. Soc.* **2005**, *127*, 2974.

[49] Bolin, K. A.; Millhauser, G. L. *Acc. Chem. Res.* **1999**, *32*, 1027.

[50] Wada, A. *Adv. Biophys.* **1976**, *9*, 1.

[51] Hol, W. G. J. *Adv. Biophys.* **1985**, *19*, 133.

[52] Richardson, J. S. *Advances in Protein Chemistry*; Academic Press, **1981**; Vol. 34.

[53] Hol, W. G. J. *Prog. Biophys. Mol. Biol.* **1985**, *45*, 149.

[54] Landschulz, W. H.; Johnson, P. F.; McKnight, S. L. *Science* **1988**, *240*, 1759; Berg, J. M. *Science* **1986**, *232*, 485.

[55] Brennan, R. G.; Matthews, B. W. *J. Biol. Chem.* **1989**, *264*, 1903.

[56] Pauling, L.; Corey, R. B. *Proc. Natl. Acad. Sci. USA* **1951**, *37*, 251.

[57] Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577.

[58] Nesloney, C. L.; Kelly, J. W. *Bioorg. Med. Chem.* **1996**, *4*, 739.

[59] Richardson, J. S. *Nature* **1977**, *268*, 495.

[60] Hol, W. G. J.; Halie, L. M.; Sander, C. *Nature* **1981**, *294*, 532.

[61] Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97.

[62] http://www.chembio.uoguelph.ca/educmat/phy456/456lec01.htm

[63] Petsko, G. A.; Ringe, D. *Protein Structure and Function*; New Science Press, **2004**.

[64] Jap, B. K.; Walian, P. J. *Q. Rev. Biophys.* **1990**, *23*, 367.

[65] Wimley, W. C. *Curr. Op. Struc. Biol.* **2003**, *13*, 404.

[66] Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S. *Chem. Rev.* **2001**, *101*, 3893.

[67] Gellman, S. H. *Acc. Chem. Res.* **1998**, *31*, 173.

[68] Matthews, M. M.; Traut, T. W. *J. Biol. Chem.* **1987**, *262*, 7232.

[69] Gojkovic, Z.; Sandrini, M. P. B.; Piskur, J. *Genetics* **2001**, *158*, 999.

[70] Schreiber, J. V.; Frackenpohl, J.; Moser, F.; Fleischmann, T.; Kohler, H. P. E.; Seebach, D. *Chembiochem* **2002**, *3*, 424.

[71] Beke, T.; Somlai, C.; Perczel, A. *J. Comput. Chem.* **2006**, *27*, 20.

[72] Aguilar, M. I.; Purcell, A. W.; Devi, R.; Lew, R.; Rossjohn, J.; Smitha, A. I.; Perlmutter, P. *Org. Biomol. Chem.* **2007**, *5*, 2884.

[73] Kovacs, J.; Ballina, R.; Rodin, R. L.; Balasubr.D; Applequi.J *J. Am. Chem. Soc.* **1965**, *87*, 119.

[74] Yuki, H.; Taketani, Y. *Polmer Lett.* **1972**, *10*, 373.

[75] Fernandezsantin, J. M.; Aymami, J.; Rodriguezgalan, A.; Munozguerra, S.; Subirana, J. A. *Nature* **1984**, *311*, 53.

[76] Tsuboi, M.; Shimanouchi, T.; Mizushima, S. *J. Am. Chem. Soc.* **1959**, *81*, 1406; Avignon, M.; Huong, P. V.; Lascombe, J.; Marraud, M.; Neel, J. *Biopolymers* **1969**, *8*, 69.

[77] Dado, G. P.; Gellman, S. H. *J. Am. Chem. Soc.* **1994**, *116*, 1054.

[78] Cheng, R. P.; Gellman, S. H.; DeGrado, W. F. *Chem. Rev.* **2001**, *101*, 3219.

[79] Gademann, K.; Hintermann, T.; Schreiber, J. V. *Curr. Med. Chem.* **1999**, *6*, 905.

[80] Gademann, K.; Jaun, B.; Seebach, D.; Perozzo, R.; Scapozza, L.; Folkers, G. *Helv. Chim. Acta* **1999**, *82*, 1.

[81] Vila, J. A.; Ripoll, D. R.; Scheraga, H. A. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 13075.

[82] Wu, Y. D.; Wang, D. P. *J. Am. Chem. Soc.* **1998**, *120*, 13485.

[83] Seebach, D.; Abele, S.; Gademann, K.; Guichard, G.; Hintermann, T.; Jaun, B.; Matthews, J. L.; Schreiber, J. V. *Helv. Chim. Acta* **1998**, *81*, 932.

[84] Seebach, D.; Abele, S.; Gademann, K.; Jaun, B. *Angew. Chem. Int. Ed.* **1999**, *38*, 1595.

[85] Seebach, D.; Overhand, M.; Kuhnle, F. N. M.; Martinoni, B.; Oberer, L.; Hommel, U.; Widmer, H. *Helv. Chim. Acta* **1996**, *79*, 913.

[86] O'Neil, K. T.; Degrado, W. F. *Science* **1990**, *250*, 646.

[87] Narita, M.; Doi, M.; Kudo, K.; Terauchi, Y. *Bull. Chem. Soc. Jpn.* **1986**, *59*, 3553.

[88] Appella, D. H.; Christianson, L. A.; Karle, I. L.; Powell, D. R.; Gellman, S. H. *J. Am. Chem. Soc.* **1996**, *118*, 13071.

[89] Barchi, J. J.; Huang, X. L.; Appella, D. H.; Christianson, L. A.; Durell, S. R.; Gellman, S. H. *J. Am. Chem. Soc.* **2000**, *122*, 2711.

[90] Appella, D. H.; Barchi, J. J.; Durell, S. R.; Gellman, S. H. *J. Am. Chem. Soc.* **1999**, *121*, 2309.

[91] Gung, B. W.; Zou, D.; Stalcup, A. M.; Cottrell, C. E. *J. Org. Chem.* **1999**, *64*, 2176.

[92] Porter, E. A.; Wang, X. F.; Schmitt, M. A.; Gellman, S. H. *Org. Lett.* **2002**, *4*, 3317.

[93] Claridge, T. D. W.; Goodman, J. M.; Moreno, A.; Angus, D.; Barker, S. F.; Taillefumier, C.; Watterson, M. P.; Fleet, G. W. J. *Tetrahedron Lett.* **2001**, *42*, 4251.

[94] Abele, S.; Seiler, P.; Seebach, D. *Helv. Chim. Acta* **1999**, *82*, 1559.

[95] Doerksen, R. J.; Chen, B.; Yuan, J.; Winkler, J. D.; Klein, M. L. *Chem. Comm.* **2003**, 2534.

[96] Gademann, K.; Hane, A.; Rueping, M.; Jaun, B.; Seebach, D. *Angew. Chem. Int. Ed.* **2003**, *42*, 1534; Winkler, J. D.; Piatnitski, E. L.; Mehlmann, J.; Kasparec, J.; Axelsen, P. H. *Angew. Chem. Int. Ed.* **2001**, *40*, 743.

[97] Seebach, D.; Gademann, K.; Schreiber, J. V.; Matthews, J. L.; Hintermann, T.; Jaun, B.; Oberer, L.; Hommel, U.; Widmer, H. *Helv. Chim. Acta* **1997**, *80*, 2033.

[98] Wu, Y. D.; Wang, D. P. *J. Am. Chem. Soc.* **1999**, *121*, 9352.

[99] Choi, S. H.; Guzei, I. A.; Gellman, S. H. *J. Am. Chem. Soc.* **2007**, *129*, 13780.

[100] Choi, S. H.; Guzei, I. A.; Spencer, L. C.; Gellman, S. H. *J. Am. Chem. Soc.* **2008**, *130*, 6544.

[101] Baldauf, C.; Gunther, R.; Hofmann, H. J. *J. Org. Chem.* **2006**, *71*, 1200.

[102] Sharma, G. V. M.; Jadhav, V. B.; Ramakrishna, K. V. S.; Jayaprakash, P.; Narsimulu, K.; Subash, V.; Kunwar, A. C. *J. Am. Chem. Soc.* **2006**, *128*, 14657.

[103] Jagadeesh, B.; Prabhakar, A.; Sarma, G. D.; Chandrasekhar, S.; Chandrashekar, G.; Reddy, M. S.; Jagannadh, B. *Chem. Comm.* **2007**, 371.

[104] Sadowsky, J. D.; Fairlie, W. D.; Hadley, E. B.; Lee, H. S.; Umezawa, N.; Nikolovska-Coleska, Z.; Wang, S. M.; Huang, D. C. S.; Tomita, Y.; Gellman, S. H. *J. Am. Chem. Soc.* **2007**, *129*, 139; Chakraborty, T. K.; Ghosh, S.; Jayaprakash, S.; Sharma, J.; Ravikanth, V.; Diwan, P. V.; Nagaraj, R.; Kunwar, A. C. *J. Org. Chem.* **2000**, *65*, 6441.

[105] Martinek, T. A.; Hetenyi, A.; Fulop, L.; Mandity, I. M.; Toth, G. K.; Dekany, I.; Fulop, F. *Angew. Chem. Int. Ed.* **2006**, *45*, 2396; Hartgerink, J. D.; Clark, T. D.; Ghadiri, M. R. *Chem. Eur. J.* **1998**, *4*, 1367; Pomerantz, W. C.; Abbott, N. L.; Gellman, S. H. *J. Am. Chem. Soc.* **2006**, *128*, 8730.

[106] Arvidsson, P. I.; Ryder, N. S.; Weiss, H. M.; Gross, G.; Kretz, O.; Woessner, R.; Seebach, D. *Chembiochem* **2003**, *4*, 1345; Epand, R. F.; Raguse, T. L.; Gellman, S. H.; Epand, R. M. *Biochemistry* **2004**, *43*, 9527.

[107] Arnold, U.; Hinderaker, M. P.; Nilsson, B. L.; Huck, B. R.; Gellman, S. H.; Raines, R. T. *J. Am. Chem. Soc.* **2002**, *124*, 8522.

[108] Horne, W. S.; Price, J. L.; Keck, J. L.; Gellman, S. H. *J. Am. Chem. Soc.* **2007**, *129*, 4178; Price, J. L.; Horne, W. S.; Gellman, S. H. *J. Am. Chem. Soc.* **2007**, *129*, 6376.

[109] Petersson, E. J.; Craig, C. J.; Daniels, D. S.; Qiu, J. X.; Schepartz, A. *J. Am. Chem. Soc.* **2007**, *129*, 5344.

[110] Petersson, E. J.; Schepartz, A. *J. Am. Chem. Soc.* **2008**, *130*, 821.

[111] Doel, M. T.; Jones, A. S.; Taylor, N. *Tetrahedron Lett.* **1969**, 2285.

[112] Yamazaki, T.; Komatsu, K.; Umemiya, H.; Hashimoto, Y.; Shudo, K.; Kagechika, H. *Tetrahedron Lett.* **1997**, *38*, 8363.

[113] Lowe, G.; Vilaivan, T. *J. Chem. Soc. Perkin Trans. 1* **1997**, 547.

[114] Bruckner, A. M.; Garcia, M.; Marsh, A.; Gellman, S. H.; Diederichsen, U. *Eur. J. Chem.* **2003**, 3555.

[115] Weiss, A.; Diederichsen, U. *Eur. J. Chem.* **2007**, 5531.

[116] Bruckner, A. M.; Chakraborty, P.; Gellman, S. H.; Diederichsen, U. *Angew. Chem. Int. Ed.* **2003**, *42*, 4395.

[117] Vilaivan, T.; Suparpprom, C.; Harnyuttanakorn, P.; Lowe, G. *Tetrahedron Lett.* **2001**, *42*, 5533.

[118] Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497.

[119] Hyrup, B.; Nielsen, P. E. *Bioorg. Med. Chem.* **1996**, *4*, 5.

[120] Nielsen, P. E. *Curr. Op. Struc. Biol.* **1999**, *9*, 353.

[121] Dueholm, K. L.; Egholm, M.; Behrens, C.; Christensen, L.; Hansen, H. F.; Vulpius, T.; Petersen, K. H.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. *J. Org. Chem.* **1994**, *59*, 5767.

[122] Christensen, L.; Fitzpatrick, R.; Gildea, B. D.; Petersen, K. H.; Hansen, H. F.; Koch, T.; Egholm, M.; Buchardt, O.; Nielsen, P.; Coull, J. M.; Berg, J. M. *J. Pept. Sci.* **1995**, *3*, 175.

[123] Demidov, V. V.; Potaman, V. N.; Frankkamenetskii, M. D.; Egholm, M.; Buchard, O.; Sonnichsen, S. H.; Nielsen, P. E. *Biochem. Pharmacol.* **1994**, *48*, 1310.

[124] Demidov, V.; Frankkamenetskii, M. D.; Egholm, M.; Buchardt, O.; Nielsen, P. E. *Nucleic Acids Res.* **1993**, *21*, 2103.

[125] Gildea, B. D.; Casey, S.; MacNeill, J.; Perry-O'Keefe, H.; Sorensen, D.; Coull, J. M. *Tetrahedron Lett.* **1998**, *39*, 7255.

[126] Sen, A.; Nielsen, P. E. *Nucleic Acids Res.* **2007**, *35*, 3367.

[127] Ishizuka, T.; Yoshida, J.; Yamamoto, Y.; Sumaoka, J.; Tedeschi, T.; Corradini, R.; Sforza, S.; Komiyama, M. *Nucleic Acids Res.* **2008**, *36*, 1464; Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem. Int. Ed.* **1996**, *35*, 1939.

[128] Egholm, M.; Behrens, C.; Christensen, L.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. *J. Chem. Soc. Chem. Comm.* **1993**, 800; Egholm, M.; Nielsen, P. E.; Buchardt, O.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 9677; Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*,1895.

[129] Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993**, *365*, 566.

[130] DeMesmaeker, A.; Altmann, K. H.; Waldner, A.; Wendeborn, S. *Curr. Op. Struc. Biol.* **1995**, *5*, 343.

[131] Rose, D. J. *Anal. Chem.* **1993**, *65*, 3545.

[132] Nielsen, P. E. *Acc. Chem. Res.* **1999**, *32*, 624.

[133] Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem. Int. Ed.* **1998**, *37*, 2797.

[134] Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Norden, B.; Graslund, A. *J. Am. Chem. Soc.* **1996**, *118*, 5544.

[135] Nielsen, P. E.; Haaima, G. *Chem. Soc. Rev.* **1997**, *26*, 73.

[136] Hyrup, B.; Egholm, M.; Nielsen, P. E.; Wittung, P.; Norden, B.; Buchardt, O. *J. Am. Chem. Soc.* **1994**, *116*, 7964.

[137] Hyrup, B.; Egholm, M.; Buchardt, O.; Nielsen, P. E. *Bioorg. Med. Chem Lett.* **1996**, *6*, 1083.

[138] Vilaivan, T.; Srisuwannaket, C. *Org. Lett.* **2006**, *8*, 1897.

[139] Jordan, S.; Schwemler, C.; Kosch, W.; Kretschmer, A.; Stropp, U.; Schwenner, E.; Mielke, B. *Bioorg. Med. Chem Lett.* **1997**, *7*, 687.

[140] Lagriffoule, P.; Wittung, P.; Eriksson, M.; Jensen, K. K.; Norden, B.; Buchardt, O.; Nielsen, P. E. *Chem. Eur. J.* **1997**, *3*, 912.

[141] Englund, E. A.; Xu, Q.; Witschi, M. A.; Appella, D. H. *J. Am. Chem. Soc.* **2006**, *128*, 16456.

[142] Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J. Am. Chem. Soc.* **2006**, *128*, 10258.

[143] Wittung, P.; Nielsen, P.; Norden, B. *J. Am. Chem. Soc.* **1996**, *118*, 7049.

[144] Betts, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science* **1995**, *270*, 1838.

[145] Datta, B.; Schmitt, C.; Armitage, B. A. *J. Am. Chem. Soc.* **2003**, *125*, 4111.

[146] Modi, S.; Wani, A. H.; Krishnan, Y. *Nucleic Acids Res.* **2006**, *34*, 4354.

[147] Datta, B.; Bier, M. E.; Roy, S.; Armitage, B. A. *J. Am. Chem. Soc.* **2005**, *127*, 4199; Krishnan-Ghosh, Y.; Stephens, E.; Balasubramanian, S. *Chem. Comm.* **2005**, 5278.

[148] DeMesmaeker, A.; Lebreton, J.; Waldner, A.; Fritsch, V.; Wolf, R. M. *Bioorg. Med. Chem Lett.* **1994**, *4*, 873.

[149] Lebreton, J.; DeMesmaeker, A.; Waldner, A.; Fritsch, V.; Wolf, R. M.; Freier, S. M. *Tetrahedron Lett.* **1993**, *34*, 6383.

[150] DeMesmaeker, A.; Lebreton, J.; Waldner, A.; Fritsch, V.; Wolf, R. M.; Freier, S. M. *Synlett* **1993**, 733.

[151] Lebreton, J.; Waldner, A.; Fritsch, V.; Wolf, R. R.; DeMesmaeker, A. *Tetrahedron Lett.* **1994**, *35*, 5225.

[152] Idziak, I.; Just, G.; Damha, M. J.; Giannaris, P. A. *Tetrahedron Lett.* **1993**, *34*, 5417.

[153] Chan, M. Y.; Fairhurst, R. A.; Collingwood, S. P.; Fisher, J.; Arnold, J. R. P.; Cosstick, R.; O'Neil, I. A. *J. Chem. Soc. Perkin Trans. 1* **1999**, 315.

[154] Rozners, E.; Katkevica, D.; Bizdena, E.; Stromberg, R. *J. Am. Chem. Soc.* **2003**, *125*, 12125.

[155] Gogoi, K.; Gunjal, A. D.; Phalgune, U. D.; Kumar, V. A. *Org. Lett.* **2007**, *9*, 2697.

[156] Gait, M. J.; Jones, A. S.; Jones, M. D.; Shepherd, M. J.; Walker, R. T. *J. Chem. Soc. Perkin Trans. 1* **1979**, 1389.

[157] Pallan, P. S.; von Matt, P.; Wilds, C. J.; Altmann, K. H.; Egli, M. *Biochemistry* **2006**, *45*, 8048.

[158] Freier, S. M.; Altmann, K. H. *Nucleic Acids Res.* **1997**, *25*, 4429.

[159] Gogoi, K.; Gunjal, A. D.; Kumar, V. A. *Chem. Comm.* **2006**, 2373.

[160] Chandrasekhar, S.; Reddy, G. P. K.; Kiran, M. U.; Nagesh, C.; Jagadeesh, B. *Tetrahedron Lett.* **2008**, *49*, 2969.

[161] Threlfall, R.; Davies, A.; Howarth, N.; Cosstick, R. *Nucleosides, Nucleotides & Nucleic Acids* **2007**, *26*, 611.

[162] Mitsuya, H.; Weinhold, K. J.; Furman, P. A.; Stclair, M. H.; Lehrman, S. N.; Gallo, R. C.; Bolognesi, D.; Barry, D. W.; Broder, S. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 7096.

[163] Pathak, T. *Chem. Rev.* **2002**, *102*, 1623.

[164] Hager, M. W.; Liotta, D. C. *J. Am. Chem. Soc.* **1991**, *113*, 5117.

[165] Hrebabecky, H.; Holy, A. *Carbohydrate Res.* **1991**, *216*, 179.

[166] Glinski, R. P.; Samikhan, M.; Kalamas, R. L.; Stevens, C. L.; Sporn, M. B. *J. Chem. Soc. D Chem. Comm.* **1970**, 915.

[167] Horwitz, J. P.; Chua, J.; Noel, M. *J. Org. Chem.* **1964**, *29*, 2076.

[168] Lin, T. S.; Prusoff, W. H. *J. Med. Chem.* **1978**, *21*, 109.

[169] Fox, J. J.; Miller, N. C. *J. Org. Chem.* **1963**, *28*, 936.

[170] Zuckermann, R.; Corey, D.; Schultz, P. *Nucleic Acids Res.* **1987**, *15*, 5305; Cosford, N. D. P.; Schinazi, R. F. *J. Org. Chem.* **1991**, *56*, 2161; Sabbagh, G.; Fettes, K. J.; Gosain, R.; O'Neil, I. A.; Cosstick, R. *Nucleic Acids Res.* **2004**, *32*, 495.

[171] Varma, R., Hogan, M., Revankar, G. R.; Int. Pat. No. WO92/18518 W.I.P.O **1992**

[172] Cole, R.; Int. Pat. No. WO98/58256 W.I.P.O **1998**

[173] Czernecki, S.; Valery, J. M. *Synthesis* **1991**, 239.

[174] Mitsunobu, O. *Synthesis* **1981**, 1.

[175] Scriven, E. F. V.; Turnbull, K. *Chem. Rev.* **1988**, *88*, 297.

[176] Hiebl, J.; Zbiral, E. *Tetrahedron Lett.* **1990**, *31*, 4007; Hiebl, J.; Zbiral, E.; Balzarini, J.; Declercq, E. *J. Med. Chem.* **1991**, *34*, 1426.

[177] Saha, A. K.; Schairer, W.; Upson, D. A. *Tetrahedron Lett.* **1993**, *34*, 8411.

[178] Zaitseva, G. V.; Kvasyuk, E. I.; Vaaks, E. V.; Barai, V. N.; Bokut, S. B.; Zinchenko, A. I.; Mikhailopulo, I. A. *Nucleosides, Nucleotides & Nucleic Acids* **1994**, *13*, 819; Zeidler, J. M.; Kim, M. H.; Watanabe, K. A. *Nucleosides, Nucleotides & Nucleic Acids* **1990**, *9*, 629.

[179] Filippov, D.; Meeuwenoord, N. J.; van der Marel, G. A.; Efimov, V. A.; KuylYeheskiely, E.; vanBoom, J. H. *Synlett* **1996**, 769; Varma, R. S.; Hogan, M. E. *Tetrahedron Lett.* **1992**, *33*, 7719.

[180] Gogoi, K.; Kumar, V. A. *Chem. Comm.* **2008**, 706.

[181] Gololobov, Y. G.; Zhmurova, I. N.; Kasukhin, L. F. *Tetrahedron* **1981**, *37*, 437.

[182] Adachi, T.; Yamada, Y.; Inoue, I. *Synthesis* **1977**, 45.

[183] Wang, G. Y.; Stoisavljevic, V. *Nucleosides, Nucleotides & Nucleic Acids* **2000**, *19*, 1413.

[184] Corey, E. J.; Nicolaou, K. C.; Balanson, R. D.; Machida, Y. *Synthesis* **1975**, 590.

[185] Barton, D. H. R.; Gero, S. D.; Quicletsire, B.; Samadi, M. *Tetrahedron Lett.* **1989**, *30*, 4969.

[186] Hutchison, A. J.; Williams, M.; Dejesus, R.; Yokoyama, R.; Oei, H. H.; Ghai, G. R.; Webb, R. L.; Zoganas, H. C.; Stone, G. A.; Jarvis, M. F. *J. Med. Chem.* **1990**, *33*, 1919.

[187] Jung, M. E.; Xu, Y. *Heterocycles* **1998**, *47*, 349.

[188] Isono, Y.; Hoshino, M.; Sudo, T. *Agricult. Biol. Chem.* **1988**, *52*, 2135.

[189] Mahmoudian, M.; Rudd, B. A. M.; Cox, B.; Drake, C. S.; Hall, R. M.; Stead, P.; Dawson, M. J.; Chandler, M.; Livermore, D. G.; Turner, N. J.; Jenkins, G. *Tetrahedron* **1998**, *54*, 8171.

[190] Pfitzner, K. E.; Moffatt, J. G. *J. Am. Chem. Soc.* **1965**, *87*, 5661.

[191] Barton, D. H. R.; Gero, S. D.; Quicletsire, B.; Samadi, M. *J. Chem. Soc. Perkin Trans. 1* **1991**, 981.

[192] Calcerrada-Munoz, N.; O'Neil, I.; Cosstick, R. *Nucleosides, Nucleotides & Nucleic Acids* **2001**, *20*, 1347.

[193] Dess, D. B.; Martin, J. C. *J. Am. Chem. Soc.* **1991**, *113*, 7277.

[194] DeMico, A.; Margarita, R.; Parlanti, L.; Vescovi, A.; Piancatelli, G. *J. Org. Chem.* **1997**, *62*, 6974.

[195] Epp, J. B.; Widlanski, T. S. *J. Org. Chem.* **1999**, *64*, 293.

196 deNooy, A. E. J.; Besemer, A. C.; vanBekkum, H. *Synthesis* **1996**, 1153.

197 Bobbitt, J. M.; Flores, M. C. L. *Heterocycles* **1988**, *27*, 509.

198 Miyazawa, T.; Endo, T.; Shiihashi, S.; Okawara, M. *J. Org. Chem.* **1985**, *50*, 1332.

199 Cella, J. A.; Kelley, J. A.; Kenehan, E. F. *J. Org. Chem.* **1975**, *40*, 1860.

200 Golubev, V. A.; Sen, V. D.; Kulyk, I. V.; Aleksandrov, A. L. *Bull. Acad. Sci. USSR. Div. Chem. Sci.* **1975**, *24*, 2119.

201 Anelli, P. L.; Biffi, C.; Montanari, F.; Quici, S. *J. Org. Chem.* **1987**, *52*, 2559.

202 Semmelhack, M. F.; Schmid, C. R.; Cortes, D. A.; Chou, C. S. *J. Am. Chem. Soc.* **1984**, *106*, 3374.

203 Bolm, C.; Magnus, A. S.; Hildebrand, J. P. *Org. Lett.* **2000**, *2*, 1173.

204 Semmelhack, M. F.; Chou, C. S.; Cortes, D. A. *J. Am. Chem. Soc.* **1983**, *105*, 4492.

205 Miyamoto, K.; Tada, N.; Ochiai, M. *J. Am. Chem. Soc.* **2007**, *129*, 2772.

206 Flyunt, R.; Theruvathu, J. A.; Leitzke, A.; von Sonntag, C. *J. Chem. Soc. Perkin Trans. 2* **2002**, 1572.

207 Sieber, P. *Tetrahedron Lett.* **1987**, *28*, 2107.

208 Kadereit, D.; Deck, P.; Heinemann, I.; Waldmann, H. *Chem. Eur. J.* **2001**, *7*, 1184.

209 Breipohl, G.; Knolle, J.; Langner, D.; Omalley, G.; Uhlmann, E. *Bioorg. Med. Chem Lett.* **1996**, *6*, 665.

210 Meienhofer, J.; Waki, M.; Heimer, E. P.; Lambros, T. J.; Makofske, R. C.; Chang, C. D. *Int. J. Pept. Protein Res.* **1979**, *13*, 35.

211 Carpino, L. A.; Mansour, E. M. E.; Knapczyk, J. *J. Org. Chem.* **1983**, *48*, 666.

212 Wade J.D.; Bedford J.; Sheppard R.C.; G.W., T. *Peptide Res.* **1991**, *4*, 194.

213 Gude, M.; Ryf, J.; White, P. D. *Lett. Pept. Sci.* **2002**, *9*, 203.

214 Taylor, C. K.; Abel, P. W.; Hulce, M.; Smith, D. D. *J. Pept. Res.* **2005**, *65*, 84; Hyde, C.; Johnson, T.; Sheppard, R. C. *J. Chem. Soc. Chem. Comm.* **1992**, 1573.

215 Narita, M.; Tomotake, Y.; Isokawa, S.; Matsuzawa, T.; Miyauchi, T. *Macromolecules* **1984**, *17*, 1903; Merrifield, R. B.; Singer, J.; Chait, B. T. *Anal. Biochem.* **1988**, *174*, 399.

216 Sarin, V. K.; Kent, S. B. H.; Merrifield, R. B. *J. Am. Chem. Soc.* **1980**, *102*, 5463.

217 Kaiser, E.; Colescot, R. I.; Bossinge, C. D.; Cook, P. I. *Anal. Biochem.* **1970**, *34*, 595.

[218] Armarego, W. L. F.; Chai, C. L. L. *Purification of Laboratory Chemicals*; 5th ed.; Butterworth-Heinemann, **2003**.

[219] Hancock, W. S.; Battersby, J. E. *Anal. Biochem.* **1976**, *71*, 260.

[220] Vojkovsky, T. *Peptide Res.* **1995**, *8*, 236.

[221] Rink, H. *Tetrahedron Lett.* **1987**, *28*, 3787.

[222] Sieber, P. *Tetrahedron Lett.* **1987**, *28*, 6147.

[223] Threlfall, R.; Davies, A.; Howarth, N. M.; Fisher, J.; Cosstick, R. *Chem. Comm.* **2008**, 585.

[224] Andersen, N. H.; Neidigh, J. W.; Harris, S. M.; Lee, G. M.; Liu, Z. H.; Tong, H. *J. Am. Chem. Soc.* **1997**, *119*, 8547.

[225] Macrae, C. F.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Shields, G. P.; Taylor, R.; Towler, M.; van De Streek, J. *J. Appl. Crystallogr.* **2006**, *39*, 453.

[226] Leitzel, J. C.; Lynn, D. G. *Chem. Rec.* **2001**, *1*, 53.

[227] Wojciechowski, F.; Hudson, R. H. E. *J. Org. Chem.* **2008**, *73*, 3807.

[228] Thomson, S. A.; Josey, J. A.; Cadilla, R.; Gaul, M. D.; Hassman, C. F.; Luzzio, M. J.; Pipe, A. J.; Reed, K. L.; Ricca, D. J.; Wiethe, R. W.; Noble, S. A. *Tetrahedron* **1995**, *51*, 6179.

[229] Bergmann, F.; Bannwarth, W.; Tam, S. *Tetrahedron Lett.* **1995**, *36*, 6823.

[230] Zaramella, S.; Stromberg, R.; Yeheskiely, E. *Eur. J. Chem.* **2002**, 2633.

[231] Challa, H.; Bruice, T. C. *Bioorg. Med. Chem.* **2004**, *12*, 1475.

[232] Nishino, S.; Yamamoto, H.; Nagato, Y.; Ishido, Y. *Nucleosides, Nucleotides & Nucleic Acids* **1986**, *5*, 159.

[233] Zheng, H. C.; Cheng, C. M.; Wang, H.; Xu, S. S.; Zhao, Y. F. *Synlett* **2004**, 2585.

[234] Meena; Sam, M.; Pierce, K.; Szostak, J. W.; McLaughlin, L. W. *Org. Lett.* **2007**, *9*, 1161; Herdewijn, P.; Balzarini, J.; Baba, M.; Pauwels, R.; Vanaerschot, A.; Janssen, G.; Declercq, E. *J. Med. Chem.* **1988**, *31*, 2040.

[235] Herdewijn, P.; Vanaerschot, A. *Tetrahedron Lett.* **1989**, *30*, 855; Battistini, C.; Giordani, A.; Ermoli, A.; Franceschi, G. *Synthesis* **1990**, 900.

[236] Gosselin, G.; Boudou, V.; Griffon, J. F.; Pavia, G.; Pierra, C.; Imbach, J. L.; Faraj, A.; Sommadossi, J. P. *Nucleosides, Nucleotides & Nucleic Acids* **1998**, *17*, 1731.

[237] Matsuda, A.; Yasuoka, J.; Sasaki, T.; Ueda, T. *J. Med. Chem.* **1991**, *34*, 999.

[238] Herdewijn, P. A. M. *J. Org. Chem.* **1988**, *53*, 5050.

[239] Yamashita, M.; Kawai, Y.; Uchida, I.; Komori, T.; Kohsaka, M.; Imanaka, H.; Sakane, K.; Setoi, H.; Teraji, T. *J. Antibiot.* **1984**, *37*, 1284.

[240] Wardrop, D. J.; Burge, M. S. *Chem. Comm.* **2004**, 1230.

[241] Yadav, J. S.; Reddy, B. V. S.; Reddy, G. *New J. Chem.* **2000**, *24*, 571.

[242] Ariza, X.; Urpi, F.; Vilarrasa, J. *Tetrahedron Lett.* **1999**, *40*, 7515.

[243] Benati, L.; Bencivenni, G.; Leardini, R.; Nanni, D.; Minozzi, M.; Spagnolo, P.; Scialpi, R.; Zanardi, G. *Org. Lett.* **2006**, *8*, 2499; Herforth, C.; Wiesner, J.; Heidler, P.; Sanderbrand, S.; Van Calenbergh, S.; Jomaa, H.; Link, A. *Bioorg. Med. Chem.* **2004**, *12*, 755.

[244] Zheng, P. W.; Duan, X. M.; Wang, W.; Wang, X. J.; Guo, Y. P.; Tu, Q. D. *Chin. J. Chem.* **2006**, *24*, 825.

[245] Nguyen-Trung, N. Q.; Botta, O.; Terenzi, S.; Strazewski, P. *J. Org. Chem.* **2003**, *68*, 2038; Van Rompaey, P.; Jacobson, K. A.; Gross, A. S.; Gao, Z. G.; Van Calenbergh, S. *Bioorg. Med. Chem.* **2005**, *13*, 973; Gao, Z. G.; Duong, H. T.; Sonina, T.; Kim, S. K.; Van Rompaey, P.; Van Calenbergh, S.; Mamedova, L.; Kim, H. O.; Kim, M. J.; Kim, A. Y.; Liang, B. T.; Jeong, L. S.; Jacobson, K. A. *J. Med. Chem.* **2006**, *49*, 2689.

[246] Holletz, T.; Cech, D. *Synthesis* **1994**, 789.

[247] Lipshutz, B. H.; Blomgren, P. A. *Org. Lett.* **2001**, *3*, 1869.

[248] Bayley, H.; Standring, D. N.; Knowles, J. R. *Tetrahedron Lett.* **1978**, 3633; Robins, M. J.; Doboszewski, B.; Nilsson, B. L.; Peterson, M. A. *Nucleosides, Nucleotides & Nucleic Acids* **2000**, *19*, 69.

[249] Chen, J. K.; Schultz, R. G.; Lloyd, D. H.; Gryaznov, S. M. *Nucleic Acids Res.* **1995**, *23*, 2661.

[250] Saran, A.; Ojha, R. P. *J. Biosci.* **1991**, *16*, 29.

[251] Vinayak, R.; van der Laan, A. C.; Brill, R.; Otteson, K.; Andrus, A.; Kuyl-Yeheskiely, E.; van Boom, J. H. *Nucleosides, Nucleotides & Nucleic Acids* **1997**, *16*, 1653.

[252] Lowe, G.; Vilaivan, T. *J. Chem. Soc. Perkin Trans. 1* **1997**, 555.

[253] Tan, T. H. S.; Hickman, D. T.; Morral, J.; Beadham, I. G.; Micklefield, J. *Chem. Comm.* **2004**, 516.

[254] Shirude, P. S.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron Lett.* **2004**, *45*, 3085.

[255] Dandapani, S. *QSAR & Comb. Sci.* **2006**, *25*, 681.

[256] Mizuno, M.; Goto, K.; Miura, T.; Inazu, T. *QSAR & Comb. Sci.* **2006**, *25*, 742.

[257] Howe-Grant, M. *Fluorine Chemistry: A Comprehensive Treatment*; Wiley VCH: New York, **1995**.

[258] Riess, J. G.; Leblanc, M. *Pure App. Chem* **1982**, *54*, 2383.

[259] Munson, M. S. B. *J. Phys. Chem.* **1964**, *68*, 796.

[260] Horvath, I. T. *Acc. Chem. Res.* **1998**, *31*, 641.

[261] Horvath, I. T.; Rabai, J. *Science* **1994**, *266*, 72.

[262] Gladysz, J. A.; Curran, D. P.; Horvath, I. T. *The Handbook of Fluorous Chemistry*; Wiley VCH: New York, **2004**.

[263] Curran, D. P. *Synlett* **2001**, 1488.

[264] Herrera, V.; de Rege, P. J. F.; Horvath, I. T.; Le Husebo, T.; Hughes, R. P. *Inorg. Chem. Comm.* **1998**, *1*, 197.

[265] Chu, Q.; Yu, M. S.; Curran, D. P. *Tetrahedron* **2007**, *63*, 9890.

[266] Dembinski, R. *Eur. J. Chem.* **2004**, 2763.

[267] Markowicz, M. W.; Dembinski, R. *Org. Lett.* **2002**, *4*, 3785.

[268] Curran, D. P.; Hadida, S.; He, M. *J. Org. Chem.* **1997**, *62*, 6714.

[269] Zhang, W.; Curran, D. P. *Tetrahedron* **2006**, *62*, 11837.

[270] Curran, D. P. *Aldrichimica Acta* **2006**, *39*, 3.

[271] Filippov, D. V.; van Zoelen, D. J.; Oldfield, S. P.; van der Marel, G. A.; Overkleeft, H. S.; Drijfhout, J. W.; van Boom, J. H. *Tetrahedron Lett.* **2002**, *43*, 7809.

[272] Montanari, V.; Kumar, K. *J. Am. Chem. Soc.* **2004**, *126*, 9528.

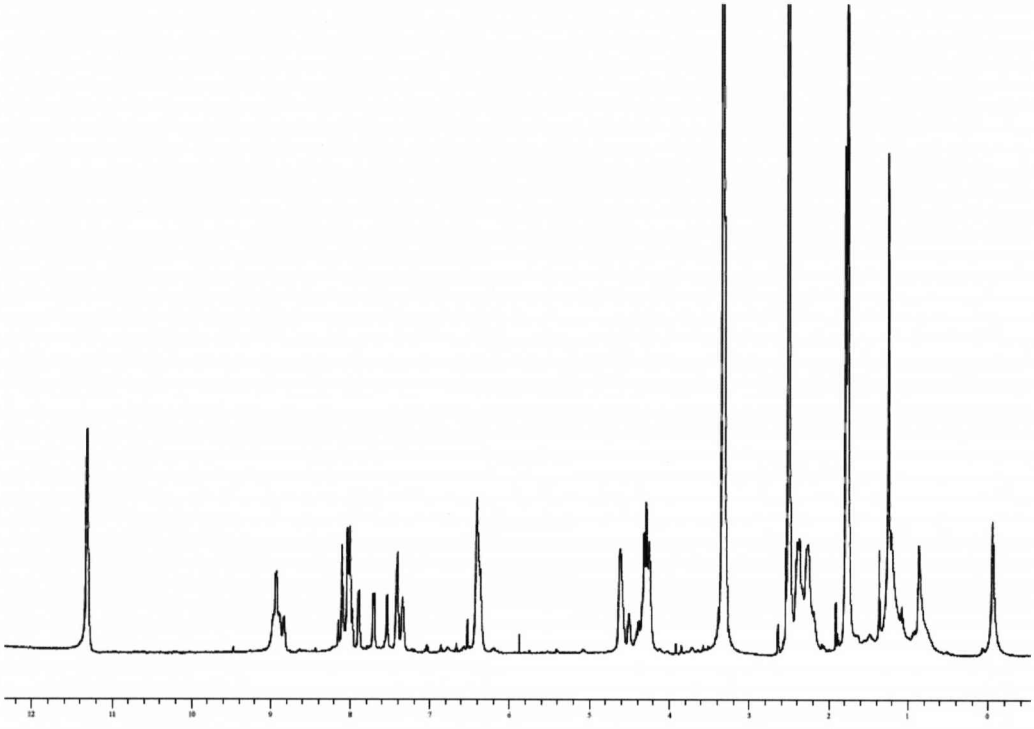
[273] ChemFiles 7, Peptide Synthesis, Sigma-Aldrich, **2007**

[274] Mizuno, M.; Goto, K.; Miura, T.; Hosaka, D.; Inazu, T. *Chem. Comm.* **2003**, 972.

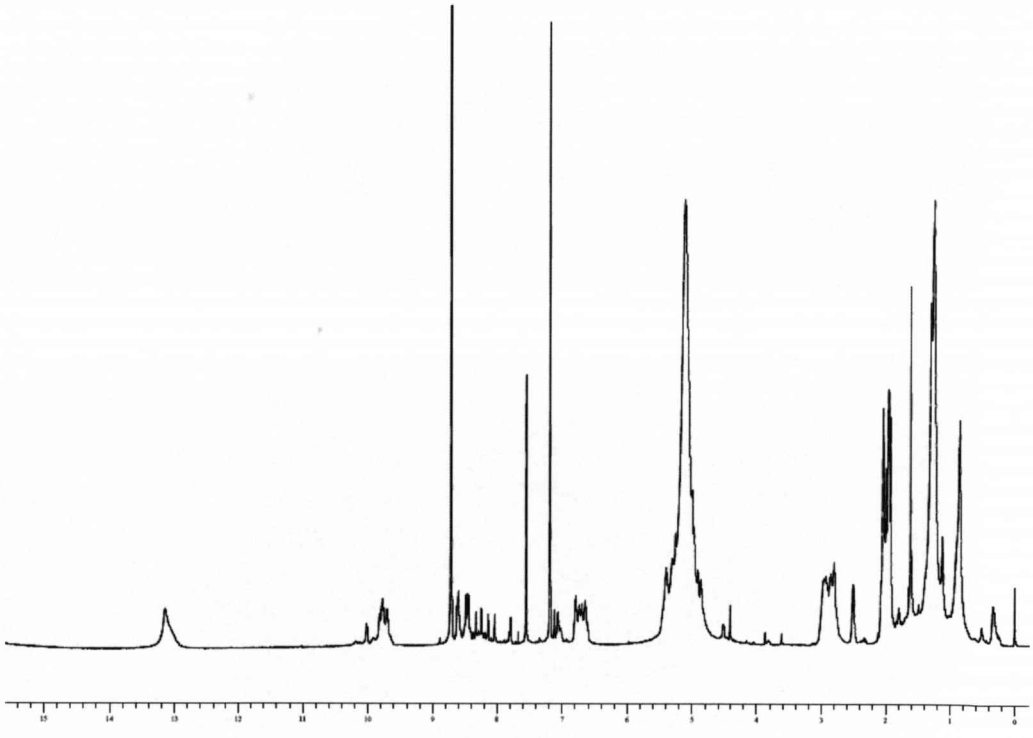
[275] Oka, N.; Shimizu, M.; Saigo, K.; Wada, T. *Tetrahedron* **2006**, *62*, 3667.

[276] Matsugi, M.; Yamanaka, K.; Inomata, I.; Takekoshi, N.; Hasegawa, M.; Curran, D. P. *QSAR & Comb. Sci.* **2006**, *25*, 713.

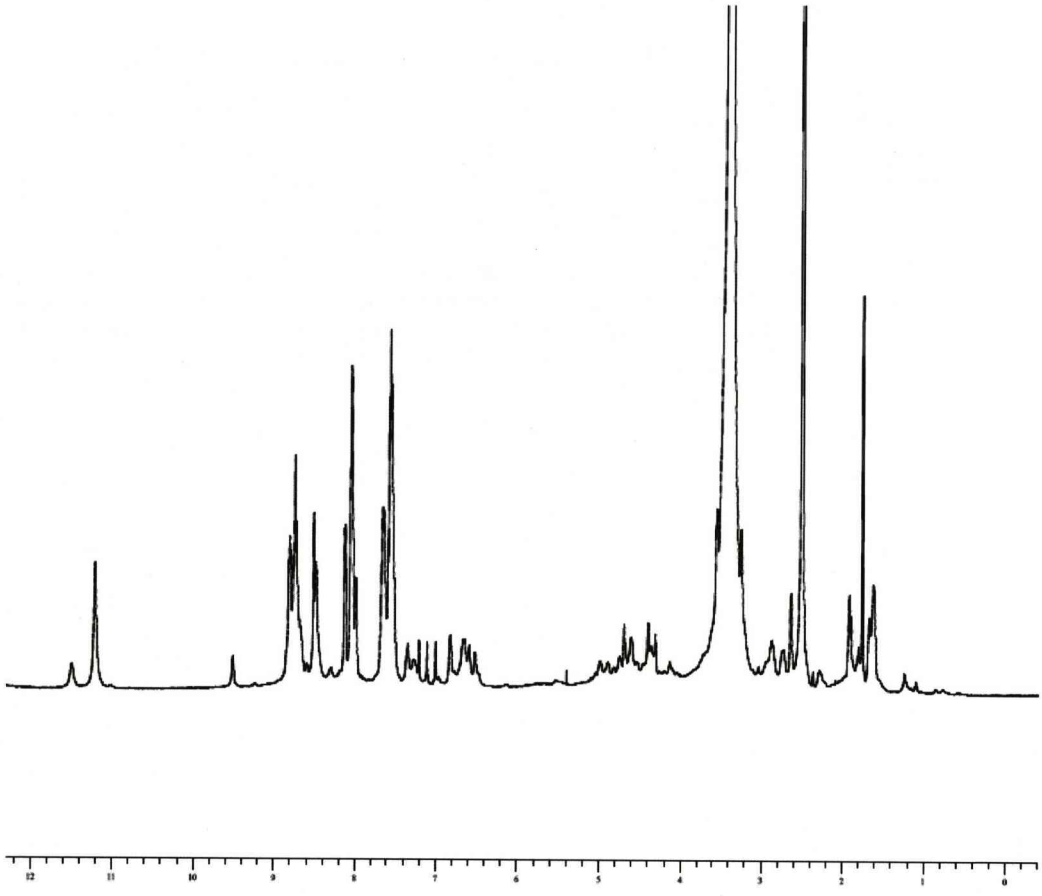
6 Appendix



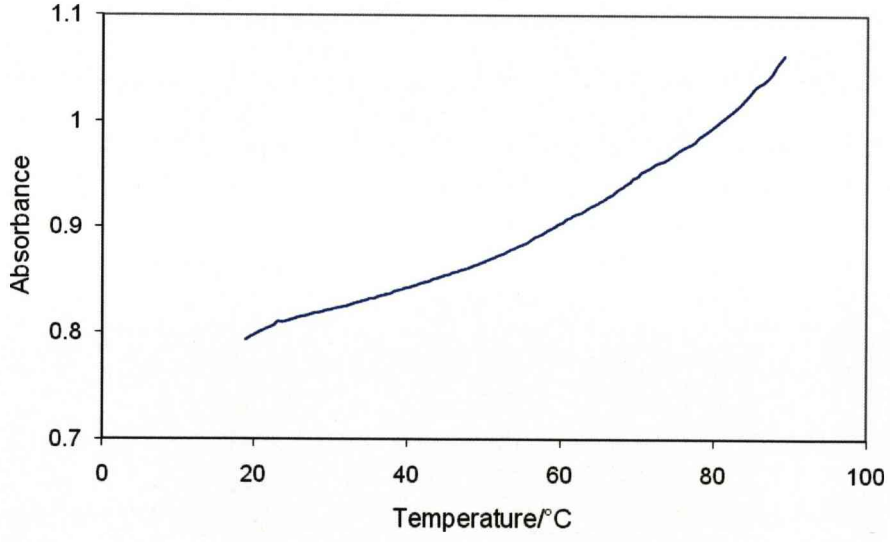
1D NMR spectrum of peptide (**12**) recorded in d_6-DMSO



1D NMR spectrum of peptide (**12**) recorded in d_5-pyridine



1D NMR spectrum of peptide (**29**) recorded in d_6-DMSO



UV melting experiment results for unmodified T_8/dA_8 duplex

Nucleosides, Nucleotides, and Nucleic Acids, 26:611–614, 2006

ISSN: 1525-7770 print / 1532-2335 online
DOI: 10.1080/15257770701490399

Taylor & Francis
Taylor & Francis Group

FOLDAMERS DERIVED FROM NUCLEOSIDE β-AMINO ACIDS: PNA OR DNA? CAN WE HAVE BOTH IN ONE PLACE?

Richard Threlfall and Andrew Davies □ *Department of Chemistry, University of Liverpool, Liverpool, United Kingdom*

Nicola Howarth □ *School of Engineering and Physical Sciences, Heriot-Watt University, Riccarton, Edinburgh, United Kingdom*

Richard Cosstick □ *Department of Chemistry, University of Liverpool, Liverpool, United Kingdom*

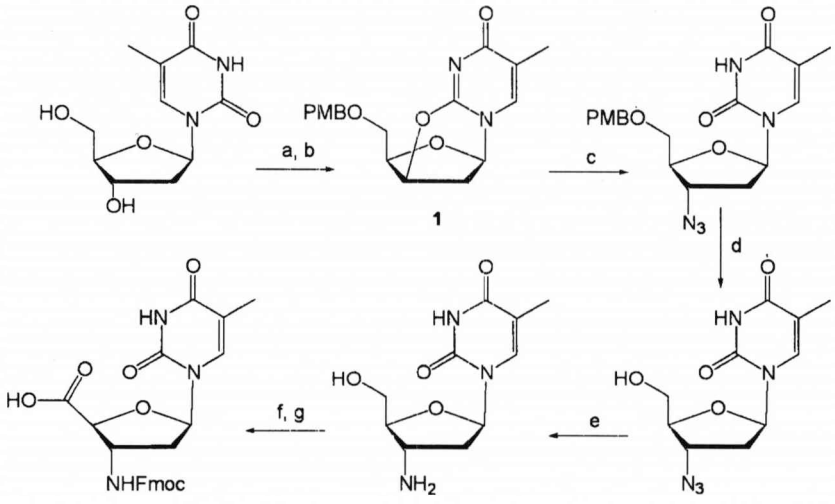
□ *The synthesis of a modified thymidine (nucleoside β-amino acid) monomer and preliminary investigations into the solid phase peptide synthesis of PNA/DNA chimeras containing a neutral, internucleoside amide linkage are described.*

Keywords Peptide nucleic acid; chimera; foldamers; nucleoside beta-amino acid

INTRODUCTION

The majority of natural biological processes rely on macromolecules with defined tertiary structures, the most obvious examples being RNA and protein catalysis. The specificity of these macromolecules is derived from their ability to adopt compact folded architectures, which, in turn, are assemblies of secondary structure motifs, such as helices and sheets. Unnatural oligomers that display similar propensity to adopt conformers are termed "foldamers".[1] Even with the emergence of peptide nucleic acid (PNA),[2] there has only been limited exploration into combining the structural potential of nucleosides (predictable nucleobase association) and the peptide bond. The synthesis of a novel nucleoside β-amino acid has been completed. Derived from thymidine, the compound contains a 5′-carboxylic acid moiety and a 3′-amino function to enable the construction of peptide bonds by standard solid-phase peptide synthesis procedures. Investigation into the synthesis of an amide-backboned oligonucleotide/polypeptide hybrid is currently underway.

Address correspondence to Richard Threlfall, Department of Chemistry, University of Liverpool, Crown Street, Liverpool L69 72D, UK. E-mail: r.threlfall@liv.ac.uk



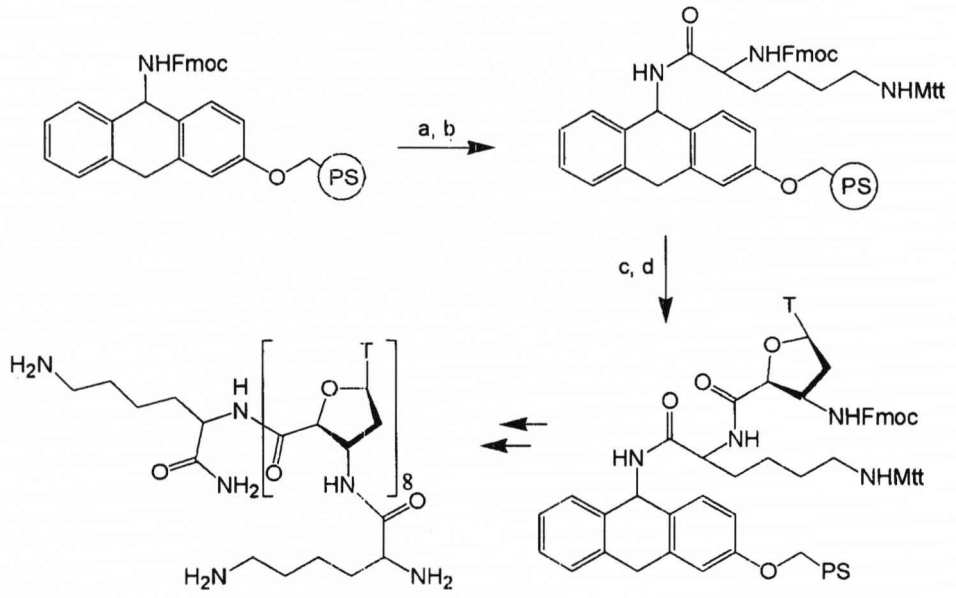
a) PPh_3, DIAD, Anisic acid, DMF; b) PPh_3, DIAD, DMF; c) LiN_3, DMF 125 °C; d) NaOMe, MeOH; e) Pd/C, H_2 MeOH; f) FmocCl, K_2CO_3, 1,4-dioxane/H_2O; g) TEMPO, BAIB, MeCN/H_2O

SCHEME 1

RESULTS AND DISCUSSION

The synthesis of the target Fmoc-protected nucleoside β-amino acid (mT, see Scheme 1) was achieved beginning with the 3′-azido-3′-deoxythymidine (AZT) synthesis of Czernecki *et al.*[3]. Starting from thymidine, two Mitsunobu style reactions performed in tandem allowed sequential, one-pot protection of 5′-OH with a PMB group whilst stereoselectively displacing 3′-OH by forming the anhydro compound (**1**) using neighboring group participation from the heterocyclic base. Yields of the anhydro species were improved by more than 10% over that obtained at room temperature by cooling the reaction mixture to a constant 15°C in a water bath during exothermic DIAD additions. The anhydro species was then opened from the α-face by N_3^- in DMF at 125°C. Little difference in yield was observed using either NaN_3 or LiN_3, both giving around 50% yield after purification. Attempts to improve the yield by the addition of a proton source (benzoic acid) to assist in the reprotonation of thymine N3 also made very little difference to the overall yield.

By far the simplest method for reduction of the 3′-azide was catalytic hydrogenation over Pd/C giving the 3′-amino compound a 97% yield. Fmoc protection of the 3′-amine was afforded by treatment of (**2**) with FmocCl and K_2CO_3 in aqueous 1,4-dioxane. From many available, the method chosen for oxidation of the 5′-OH was using free radical TEMPO and hypervalent iodine complex (bis-acetoxy)-iodobenzene (BAIB).[4] In the standard



a) 2 % DBU/DMAc; b) Fmoc-Lys-(Mtt)-OH, HBTU, DIPEA, DMAc;
c) 2 % DBU/DMAc, Compound (**3**), HBTU, DIPEA, DMAc

SCHEME 2

solvent for this reaction (aqueous acetonitrile) the Fmoc-protected nucleoside is only sparingly soluble, leading to reaction times of greater than 24 hours and low yields of around 45%.

It was believed that these neutral-backboned oligomers would not be readily soluble in aqueous media and that natural α-amino acids should be included in an attempt to improve aqueous solubility. Therefore, it was decided that lysine residues should be included at both *C*- and *N*-terminals (see Scheme 2). Initial studies on oligomerisation of the Fmoc-protected nucleoside β-amino acid were based on the method of Gude *et al.*[5] Loading of Fmoc-Lys-(monomethyltrityl)-OH onto the solid support (Sieber amide resin) was determined by UV assay of dibenzofulvene at $\lambda = 294$ and 304 nm liberated during the deblocking step.

The technique was also extended to the first two couplings of mT to inspect the coupling efficiencies with the natural amino acid and with itself. This qualitative method was sufficient to show that little difference was observed for the coupling of lysine to mT when using either Benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP) or 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HBTU) as the coupling reagent. However, NMP and DMF were particulary poor solvents. The mT to mT couplings were also solvent dependent. In this case, coupling yields were significantly reduced using the

combination of HBTU with DMF but increased greatly with the addition of 20% DMSO. The superior solvent/coupling reagent combination was found to be *N*,*N*,-dimethylacetamide with HBTU.

CONCLUSIONS

The modified thymidine monomer, mT, has been synthesized successfully in reasonable yield from thymidine and the production of oligomers via solid-phase peptide synthesis protocols is currently under investigation.

REFERENCES

1. Gellman, S.H. Foldamers: A manifesto. *Accounts Chem. Res.* **1998**, 31, 173–180.
2. Nielsen, P.E.; Egholm, M.; Berg, R.H.; Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **1991**, 254, 1497–1500.
3. Czernecki, S.; Valery, J.M. An efficient synthesis of 3′-Azido-3′-deoxythymidine (AZT). *Synthesis* **1991**, 239–240.
4. Epp, J.B.; Widlanski, T.S. Facile preparation of nucleoside-5′-carboxylic acids. *J. Org. Chem.* **1999**, 64, 293–295.
5. Gude, M.; Ryf, J.; White, P.D. An accurate method for the quantitation of Fmoc-derivatized solid phase supports. *Lett. Pept. Sci.* **2002**, 9, 203–206.

Peptides derived from nucleoside β-amino acids form an unusual 8-helix†

Richard Threlfall,[a] Andrew Davies,[a] Nicola M. Howarth,[b] Julie Fisher*[c] and Richard Cosstick*[a]

Received (in Cambridge, UK) 17th September 2007, Accepted 5th November 2007
First published as an Advance Article on the web 15th November 2007
DOI: 10.1039/b714350h

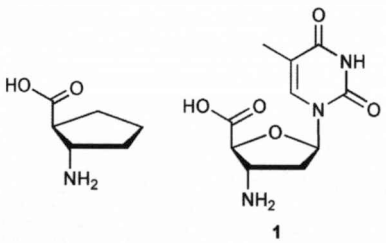
Peptides of varying length (dimers to octamers) were prepared from nucleoside β-amino acids and conformational studies, based on NOE observations, show that the β-peptides form an unusual 8-helix.

Unnatural oligomers that have a propensity to adopt well-defined compact conformations have been termed foldamers[1] and possess the potential to mimic structural motifs found in complex biopolymers such as proteins and RNA. A number of exciting applications can be envisaged for foldamers including peptidomimetics,[2] antimicrobial agents[3] and components in nanostructured materials.[4] Of all the unnatural oligomers under investigation, the folding preferences of β-peptides have been the most intensively studied. In particular, studies on β-amino acids derived from conformationally restrained systems such as cycloalkanes[5,6] and sugars,[7] show that they form peptides that adopt strongly helical structures. Gellman and co-workers have shown that β-peptides derived from *trans*-2-aminocyclopentanecarboxylic acid[5] (*trans*-ACPC, Fig. 1) adopt a 12-helix, whereas peptides prepared from *trans*-2-aminocyclohexanecarboxylic acid[6] (*trans*-ACHC), form a 14-helix.

Whilst inspiration for the construction of foldamers has largely been drawn from protein chemistry, nucleosides also possess a number of features that make them attractive monomers for this purpose. The most notable of which are well established methods for oligomer synthesis and predictive associative properties of the nucleobases. The association of peptide helices through base-pairing provides a potentially powerful method for controlling the assembly of complex tertiary structures. As proof of principle, Brückner *et al.* have demonstrated that helices containing β-homoalanine functionalised with nucleobases can associate through base pairing.[8] We were struck by the fact that nucleosides can be readily converted to β-amino acids which are closely related to Gellman's helix-forming *trans*-ACPC monomer and now report the synthesis of peptides derived from the thymidine β-amino acid **1** (Fig. 1) and their structural properties.

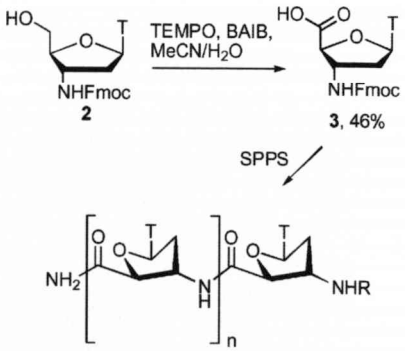
The target monomer[9] **3** (Scheme 1), was synthesised from the previously reported[10] fluorenylmethoxycarbonyl (Fmoc)-protected amino alcohol **2** by oxidation with [bis(acetoxy)iodo]benzene (BAIB) and 2,2,6,6-tetramethypiperidinyloxy (TEMPO)[11] in aqueous acetonitrile. Although the oxidation proceeds in relatively low yield (46%) the method has the advantage that the Fmoc-protected amino acid precipitates out of the reaction mixture and is obtained pure after filtration and washing with diethyl ether.

The choice of resin for Fmoc solid-phase peptide synthesis (SPPS) was influenced by the sensitivity of the glycosidic bond to acid. With this in mind, the hyperacid-labile Sieber amide resin[12] was chosen, which, on cleavage, produces peptides with a primary amide at the carboxylate terminus. The peptide assembly used 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) in *N*,*N*-dimethylacetamide as the coupling agent and followed standard Fmoc protocols.[13] Using this procedure, dimers, tetramers and octamers, both with and without the Fmoc group, were prepared (Scheme 1). Following cleavage from the support and washing with water, the oligomers gave clean NMR spectra and were also shown, by mass spectrometry, to be free of failure sequences.



1

Fig. 1 *trans*-ACPC and thymidine β-amino acid (**1**).



2

3, 46%

4 R = H, n = 1
5 R = H, n = 3
6 R = Fmoc, n = 3
7 R = H, n = 7
8 R = Fmoc, n = 7

Scheme 1 T = thymin-1-yl.

[a]*Department of Chemistry, University of Liverpool, Crown Street, Liverpool, UK L69 7ZD. E-mail: rcosstic@liv.ac.uk; Fax: +44 (0) 151 794 3588; Tel: +44 (0) 151 794 3514*
[b]*Chemistry, School of Engineering & Physical Sciences, William H. Perkin Building, Heriot Watt University, Edinburgh, UK EH14 4AS*
[c]*University of Leeds, Leeds, UK LS2 9JT. E-mail: J.Fisher@chem.leeds.ac.uk; Fax: +44 (0) 113 343 6565; Tel: +44 (0) 113 343 6577*

† Electronic supplementary information (ESI) available: Experimental details and NMR spectra. See DOI: 10.1039/b714350h

Table 1 Temperature-dependent chemical shifts of NH protons of **6** in d_6-DMSO solution

Residue number	25 °C	30 °C	35 °C	40 °C	45 °C	Average shift difference/ppb K^{-1}
1	8.845	8.818	8.790	8.764	8.729	5.80
2	8.910	8.882	8.857	8.830	8.783	6.35
3	8.980	8.951	8.926	8.902	8.851	6.45
4	8.117	8.098	8.084	8.069	8.054	3.15

Conformational features of the Fmoc-protected tetramer **6** and the fully deprotected octamer **7**, in d_6-DMSO and d_5-pyridine solution respectively, have been determined using NMR spectroscopy.

A complete ^{1}H resonance assignment has been achieved for **6** through the combined use of DQF-COSY, TOCSY and NOESY spectra (see ESI for numbering and NMR assignment details†). The amide NHs are well resolved and through their chemical shifts in the 2D spectra the remaining assignments could be made. The NH chemical shifts are worthy of note. The NH of residue 4 (the Fmoc-protected amino terminus) is almost identical to that observed for the Fmoc-protected monomer. This observation, together with the lowfield shift of the remaining secondary amide protons, is supportive of a hydrogen bonding network involving the three interresidue NHs.[14] Variable temperature experiments performed over the temperature range 25 to 45 °C resulted in temperature dependent chemical shift changes to highfield of approximately 6 ppb K^{-1} for the NHs of residues 1 to 3, but only 3 ppb K^{-1} for residue 4 (Table 1).

It is the convention to assume that small shift changes would arise for amide protons involved in intramolecular hydrogen bonds and large changes for solvent exposed NHs. However, for essentially linear systems, as here, if hydrogen bonding is intramolecular then a large temperature coefficient would be observed.[15] Thus these results are also consistent with hydrogen bonding within the tetramer. The possibility of a hydrogen bonding array is also suggested by the model, formed through consideration of a series of NOE connectivities identified in the NOESY spectrum, and built using Macro-Model.[16] More specifically, through-space connections are observed between NH(1)–H3′(2), 3H6–2H3′, NH(3)–H4′(4). These, together with a number of other interresidue NOEs and intraresidue connections (see Fig. 2 and ESI†), have been utilised to build a model for the tetramer (Fig. 3). All of the NOEs observed are consistent with this model. NMR studies of **4** had revealed the sugar puckers of the first and second rings to be 60% and 50% south, respectively. Consequently, in the modeling procedure sugar puckers were constrained to 50% south, with dihedral angles permitted to vary by $\pm 10°$.

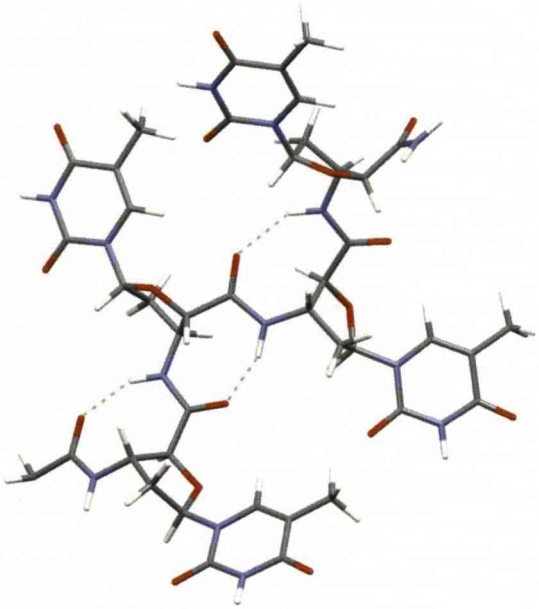


Fig. 3 Model of the structure formed by tetramer **6** produced using Macro-Model and NMR constraints.[18,19] The dashed lines are provided to indicate hydrogen bonded NHs and carbonyls. The Fmoc group has been deleted for ease of viewing.

In Fig. 3 a zig-zag arrangement of the deoxyribose rings is apparent, leading to the thymine bases occupying opposite faces of the extended chain. In this conformation the NH of residue 1 comes close in space to the carbonyl attached 3′ to residue 2. Similarly the NH of residue 2 comes close in space to the carbonyl 3′ to residue 3 and finally the NH of residue 3 is close in space to the carbonyl of the Fmoc group. An atom count between these groups reveals an 8-helix. Whilst 12- and 14-helices are much more common, a number of examples of 8-helix have appeared in the literature.[14] Doerksen *et al.* found the 8-helix to be energetically most favourable for their conformationally-constrained oxanorbornene β-peptides. Whilst the current systems have more flexibility in the pentose ring compared to the oxanorbornene systems, the helical arrangement observed here may be due to steric factors imposed by the thymine nucleobase. In this respect, it has been suggested that *trans*-ACPC monomers with relatively small substituents on the 4 position (which is comparable to the anomeric position of **1**) perturb the 12-helix conformation.[17]

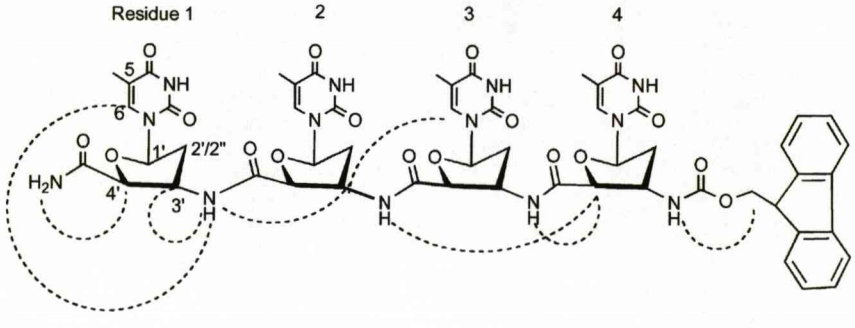


Fig. 2 Diagram of peptide **6** presenting some of the NOE connections used in determining the model for **6** shown in Fig. 3.

Similar studies have been performed for the deprotected octamer **7**, in d_5-pyridine solution, although resonance assignment is not yet complete.[18] Residues 1 and 8 have been unambiguously identified and despite the change in solvent, these look remarkably similar to the first and last residues of the tetramer **6**. A series of interresidue NH–H3′/H4′ NOE connections have been observed, together with a number of NH–NH, and NH–H6 (interresidue); the latter NOEs being more apparent at low temperature (10 °C). These NOE connectivities bear similarities to those observed for tetramer **6** and imply a similar helical arrangement for the octamer **7**, but more extensive resonance assignment is required for this to be confirmed. Variable temperature experiments over the range 20 to 45 °C have revealed chemical shift changes for the secondary amide protons of between 8.3 and 9.4 ppb K^{-1} to highfield. In contrast, the average chemical shift change for the thymine H6 protons was 3 ppb K^{-1}. These data are consistent with a hydrogen bonded system in which there is little or no base stacking (as suggested by Fig. 3). With the aid of the model derived for the tetramer **6** it is anticipated that full assignment and structure determination for the octamer **7** will be forthcoming. Similar features are observed for **7** in DMSO solution, however there is better resolution in the H6 chemical shifts for **7** in d_5-pyridine.

In conclusion, peptides derived from a thymidine β-amino acid have been prepared and their conformation studied by NMR spectroscopy. Interestingly, the tetramer **6** forms an 8-helical conformation, which differs considerably from the 12-helix derived from the related *trans*-ACPC. An important long term aim is to investigate association between helices derived from adenine and thymine and thus capable of nucleobase-pairing. These systems would offer a new structural motif which would have applications in the assembly of innovative macromolecular architectures.

We wish to thank the EPSRC and BBSRC for providing support for RT and AD, respectively. We also thank A. Mills (Liverpool) for mass spectrometry and C. Kilner (Leeds) for preparation of Fig. 3.

Notes and references

1 R. P. Cheng, S. H. Gellman and W. F. DeGrado, *Chem. Rev.*, 2001, **101**, 3219; D. Seebach, A. K. Beck and D. J. Bierbaum, *Chem. Biodiversity*, 2004, **1**, 1111; S. H. Gellman, *Acc. Chem. Res.*, 1998, **31**, 173.

2 T. K. Chakraborty, S. Ghosh, S. Jayaprakash, J. A. R. P. Sharma, V. Ravikanth, P. V. Diwan, R. Nagaraj and A. C. Kunwar, *J. Org. Chem.*, 2000, **65**, 6441; J. D. Sadowsky, W. D. Fairlie, E. B. Hadley, H. S. Lee, N. Umezawa, Z. Nikolovska-Coleska, S. M. Wang, D. C. S. Huang, Y. Tomita and S. H. Gellman, *J. Am. Chem. Soc.*, 2007, **129**, 139.

3 R. F. Epand, T. L. Raguse and S. H. Gellman, *Biochemistry*, 2004, **43**, 9527; P. I. Arvidsson, N. S. Ryder, H. M. Weiss, G. Gross, O. Kretz, R. Woessner and D. Seebach, *ChemBioChem*, 2003, **4**, 1345.

4 T. A. Martinek, A. Hetenyi, L. Fulop, I. M. Mandity, G. K. Toth, I. Dekany and F. Fulop, *Angew. Chem., Int. Ed.*, 2006, **45**, 2396.

5 J. Applequist, K. A. Bode, D. H. Appella, L. A. Christianson and S. H. Gellman, *J. Am. Chem. Soc.*, 1998, **120**, 4891.

6 D. H. Appella, L. A. Christianson, I. L. Karle, D. R. Powell and S. H. Gellman, *J. Am. Chem. Soc.*, 1996, **118**, 13071; D. H. Appella, J. J. Barchi, S. R. Durell and S. H. Gellman, *J. Am. Chem. Soc.*, 1999, **121**, 2309.

7 S. F. Barker, D. Angus, C. Taillefumier, M. R. Probert, D. J. Watkin, M. P. Watterson, T. D. W. Claridge, N. L. Hungerford and G. W. J. Fleet, *Tetrahedron Lett.*, 2001, **42**, 4247; G. V. M. Sharma, V. Subash, K. Narsimulu, A. R. Sankar and A. J. Kunwar, *Angew. Chem., Int. Ed.*, 2006, **45**, 8207; T. D. W. Claridge, J. M. Goodman, A. Moreno, D. Angus, S. F. Barker, C. Taillefumier, M. P. Watterson and G. W. J. Fleet, *Tetrahedron Lett.*, 2001, **42**, 4251; Y. Suhara, J. E. K. Hildreth and Y. Ichikawa, *Tetrahedron Lett.*, 1996, **37**, 1575.

8 A. M. Brückner, M. Garcia, A. Marsh, S. H. Gellman and U. Diederichsen, *Eur. J. Org. Chem.*, 2003, 3555; A. M. Brückner, P. Chakraborty, S. H. Gellman and U. Diederichsen, *Angew. Chem., Int. Ed.*, 2003, **42**, 4395.

9 Some preliminary work on the Fmoc-protected thymidine amino acid (**3**) has been presented: International Roundtable on Nucleosides, Nucleotides and Nucleic Acids, Bern, September 3–7, 2006. R. Threlfall, A. Davies, R. Cosstick and N. Howarth, *Nucleosides, Nucleotides Nucleic Acids*, 2007, **26**, 611.

10 N. Mignet and S. M. Gryaznov, *Nucleic Acids Res.*, 1998, **26**, 431.

11 J. B. Epp and T. S. Widlanski, *J. Org. Chem.*, 1999, **64**, 293.

12 S. Sieber, *Tetrahedron Lett.*, 1987, **28**, 2107.

13 W. C. Chan and P. D. White, *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*, Oxford University Press, New York, 2000, pp. 1–75.

14 R. J. Doerksen, B. Chen, J. Yuan, J. D. Winkler and M. L. Klein, *Chem. Commun.*, 2003, 2534, and references therein.

15 N. H. Anderson, J. W. Neidigh, S. M. Harris, G. M. Lee, Z. Liu and H. Tong, *J. Am. Chem. Soc.*, 1997, **119**, 8547.

16 T. A. Halgren, *J. Comput. Chem.*, 1999, **20**, 720.

17 T. J. Peelen, Y. Chi, E. P. English and S. H. Gellman, *Org. Lett.*, 2004, **6**, 4411.

18 Expansions of TOCSY and NOESY spectra for both **6** and **7** are included in the ESI†.

19 C. F. Macrae, P. R. Edgington, P. McCabe, E. Pidcock, G. P. Shields, R. Taylor, M. Towler and J. van de Streek, *J. Appl. Crystallogr.*, 2006, **39**, 453.

Symposium date: 8th September 2008

Nucleic Acids Symposium Series No. 52 337–338
doi:10.1093/nass/nrn170

Towards Fluorous-Phase Synthesis of Nucleoside β-Peptides

Richard Threlfall*[1], Richard Cosstick[1] and Takeshi Wada[2]

[1]Dept. of Chemistry, University of Liverpool, Crown Street, Liverpool, L69 7ZD, UK, [2]Dept. of Medical Genome Sciences, Graduate School of Frontier Sciences, University of Tokyo, Kashiwa 277-8562, Japan.

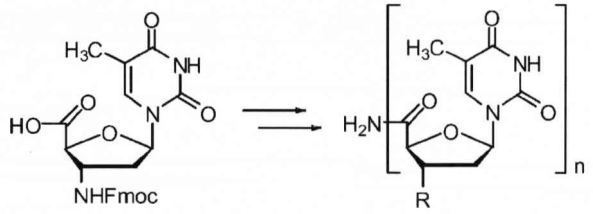
ABSTRACT

We have recently shown that peptides derived from nucleoside β-amino acids adopt an unusual 8-helical conformation in solution[1]. These β-peptide homo-oligomers were constructed using Fmoc solid-phase peptide synthesis protocols (SPPS); however, we found these procedures to be somewhat limiting in terms of scale and lacking a quantitative method of monitoring reactions. Preliminary investigations into the use of fluorous dendrons as an alternative to SPPS have been conducted and show that coupling reactions with modified nucleoside (1) and fluorous dendrimer (2) are possible in the presence of a glycine linker. Significantly, success of the coupling reactions could be monitored by NMR without loss of materials or termination of the synthetic sequence and this promises much for the future of both oligonucleotide and polypeptide synthesis.

INTRODUCTION

In many respects the development of oligonucleotide and polypeptide chemistry has followed a similar path, no more so than with the introduction of solid-phase synthesis[2,3]. Solid-phase techniques attempt to tackle the problem faced by all chemists when performing reactions in solution: separation of products and reagents. In solid-phase synthesis a substrate is attached to a surface, for example a controlled pore glass bead or a cross linked polystyrene resin, and reagents flooded over this surface in order to perform reactions. Separation is then simply a matter of washing the spent reagents from the surface and oligomeric compounds can be produced by iteration of this process.

During optimisation of solid-phase procedures *en route* to oligomers derived from modified thymidine (**1, Fig. 1**) the limitations of solid-phase synthesis in terms of scale and judging the success of coupling reactions became very apparent. In order to prevent aggregation of the growing chains during SPPS it was necessary to load the resin at a significantly lower concentration than full capacity and this has obvious implications for large-scale production. Secondly, although reaction success or failure can be assessed qualitatively by colourmetric methods[4], these are not always reliable and there are few alternative, quantitative analytical techniques available that can be performed without termination of the synthesis. It is clear that a solution-phase technique which can be carried out on a much larger scale than solid-phase methods but with equivalent purification performance is desirable. The possibility of including reaction monitoring in any such process would also be extremely attractive.

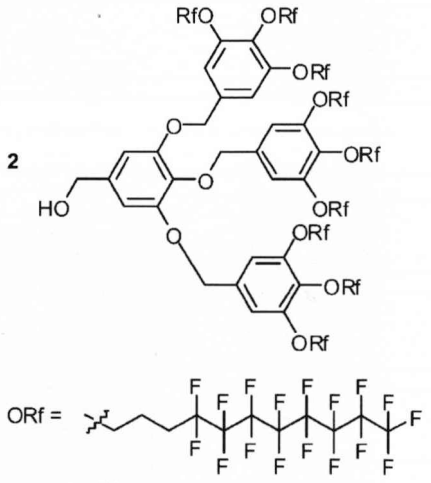


1

Fig. 1 Nucleoside β-amino acid derived from thymidine (Fmoc = 9-fluorenylmethoxycarbonyl)

Fluorous-phase chemistry has recently attracted much attention for its interesting applications in the purification of organic compounds[5]. Fluorous technology allows compounds and intermediates appended with perfluorinated tags or "fluorous supports" to be purified by simple extraction with fluorous solvents without the need for chromatography. It occurred to us that oligomers could be synthesised by a cycle of reaction followed by extraction and a major advantage of this system would be the ability to accurately monitor the success of individual couplings by NMR without cleaving the oligomer from the support.

We now present the results of initial testing of such fluorous techniques for synthesis of homo-oligomers derived from nucleoside β-amino acid (**1**) using a fluorous support (**2**, designated FD-O (H)) in place of a solid-phase resin.



2

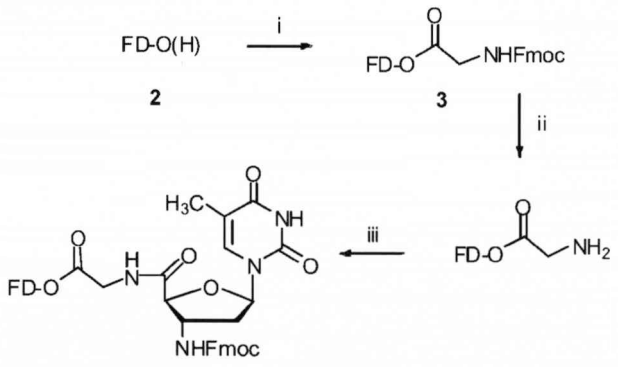
Fig. 2 Fluorous dendrimer (FD-O (H))

RESULTS AND DISCUSSION

Nucleoside β-amino acid (**1**) was synthesised according to the published route[1] and the first step in mimicking the cycle of SPPS previously applied in the production of nucleoside β-peptides involved esterification of this amino acid with fluorous dendrimer (**2**) in a mixture of miscible organic and fluorous solvents, such as DMF/HFE-7100, utilising a coupling reagent such as HBTU or NMTP[6].

Unfortunately, after screening many permutations of temperature, coupling reagents and solvents no suitable reaction conditions could be found. The failure of the coupling is thought to have been due to extensive steric interference from both nucleoside and dendrimer alike. In an effort to circumvent this problem, analogy was drawn with SPPS procedures in which the amino acid is not coupled directly with the resin, but to a linker.

It was hypothesised that a suitable linker would be sterically undemanding in order to facilitate initial reaction with the dendrimer. From previous SPPS studies (**1**) was known to successfully form amide bonds with the xanthenyl amine linker of Sieber amide resin. Thus inclusion of a much smaller amino acid such as glycine to act as a linker would decrease the steric barrier encountered by the nucleoside amino acid whilst simultaneously changing the connection to the dendrimer from ester to amide and improve the probability of successful coupling.



Scheme 1 i) Fmoc-Gly-OH, 2,4,6-Collidine, NMTP, THF ii) FC-72, 2 % DBU/DMF iii) (**1**), 2,4,6-Collidine, HBTU, THF

Coupling of the dendrimer with Fmoc-glycine-OH proceeded smoothly in under one hour using 5 equivalents of 2,4,6-collidine and 2.4 equivalents of NMTP in THF at 45°C (**scheme 1**). The reaction work-up simply consisted of addition of FC-72 (perfluorinated hexane), separating and washing the fluorous layer with chloroform then water and evaporating the FC-72 to afford coupled product (**3**) in excellent yield with no chromatography required.

As well as the synthesis being performed in solution without the need for laborious purification another and possibly more significant advantage of this fluorous process was then vividly illustrated. A sample of the reaction residue was dissolved in a mixture of $CDCl_3/C_6F_6$, analysed by NMR, then the solvents removed once more and the material recovered. The residue was successfully characterised as (**3**) by NMR and no loss of product or termination of the synthesis was necessary. Later MALDI-TOF mass spectrometry (MS) confirmed the identity of (**3**).

Removal of the Fmoc group was performed in a biphasic mix of FC-72 and 2 % DBU/DMF with a similar extractive work-up. Deprotection was observed to occur in a few minutes; unfortunately considerable cleavage of the glycine from the dendrimer was also observed. This is most likely due to formation of an oxazolone derivative by the imidate oxygen generated by deprotonation of the Fmoc amide. To suppress this side reaction deprotection should be performed with a weaker base such as piperidine, to prevent amide deprotonation. Coupling of the nucleoside β-amino acid (**1**) to the free amine of glycine was attempted with HBTU in the presence of excess collidine in THF at ambient temperature. After 24 hours an estimated 3:1 ratio of product to starting material was evident by NMR and whilst this reaction was incomplete, it was a vastly improved result compared to couplings without the glycine linker.

CONCLUSION

Although the fluorous dendrimer was reluctant to react with bulky nucleoside derivative (**1**) the reaction with α-glycine was smooth and high yielding, requiring no purification other an extractive work-up. Deprotection and further reaction of the glycine functionalised dendrimer with the nucleoside amino acid was also achieved to a lesser degree of success; however, integration of NMR monitoring into the process clearly demonstrates that fluorous-phase chemistry is a potentially powerful alternative to solid-phase oligomer synthesis.

We would like to thank the Daiwa Anglo-Japanese Foundation and JSPS for their generous financial support.

REFERENCES

1. Threlfall, R., Davies, A., Howarth, N., Fisher, J., Cosstick, R. (2008) *Chem. Commun.*, 585-587.
2. Merrifield, R.B., (1963) *J. Am. Chem. Soc.* **85**, 2149-2154
3. Matteucci, M. D., Caruthers, M. H., (1981) *J. Am. Chem. Soc.* **103**, 3185-3191.
4. Kaiser, E., Colescot, R. I., Bossinge, C.D., Cook, P. I., (1970) *Anal. Biochem.* **34**, 595-598.
5. Gladysz, J. A., Curran, D. P., Horváth, I. T., (2004) *Handbook of Fluorous Chemistry*, Wiley VCH.
6. Oka, N., Shimizu, M., Saigo, K., Wada, T., (2006) *Tetrahedron* **62**, 3667-3673.

*Corresponding Author. E-mail: r.threlfall@liv.ac.uk

Symposium date: 8th September 2008

Nucleic Acids Symposium Series No. 52 313–314
doi:10.1093/nass/nrn158

Foldamers derived from nucleoside beta-amino acids: a new twist on the DNA helix

Richard Cosstick[1]*, Richard Threlfall[1] and Julie Fisher[2]

[1] Department of Chemistry, University of Liverpool, Crown Street, Liverpool, L69 7ZD, UK and [2] School of Chemistry, University of Leeds, Leeds, LS2 9JT, UK.

ABSTRACT

Peptides derived from a thymidine beta-amino acid have been prepared by solid-phase synthesis and their conformation investigated by NMR. Interestingly, NMR and modelling studies indicate that the tetramer and octamer form an unusual 8-helical conformation. Studies are currently underway to investigate the synthesis of peptides derived from the other deoxyribonucleosides with the intention of examining the association between helices capable of nucleobase-pairing.

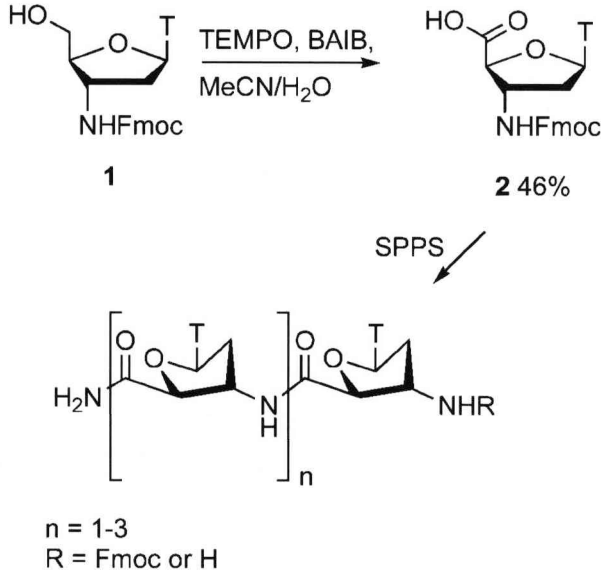
INTRODUCTION

There has been considerable interest in nucleic acid analogues in which the phosphodiester backbone is replaced by an alternative isosteric backbone. These oligomers are often designed to bind strongly to complementary RNA sequences and thus have application as antisense or RNAi agents. In general, the inherent conformational properties of single stranded nucleic acid analogues has not been an area of interest. However, the work of Gellman, has drawn attention to the potential uses of unnatural oligomers (termed foldamers) that have a propensity to adopt well-defined compact conformations, particularly helices and can thus mimic structural motifs found in complex biopolymers such as proteins and RNA.[1] Whilst inspiration for the construction of foldamers has largely been drawn from protein chemistry, nucleosides also possess a number features which make them attractive monomers for this purpose. The most notable of which are well established methods for oligomer synthesis and predictive associative properties of the nucleobases.

We were struck by the fact that nucleosides can be readily converted to β-amino acids which are closely related to Gellman's helix-forming *trans*-2-aminocyclopentanecarboxylic acid monomer[2] and now report the synthesis of peptides derived from the thymidine beta-amino acid (**Scheme 1**) and their structural properties.[3] The analogous amino acid derived from 2'-deoxyadenosine has also been prepared.

RESULTS AND DISCUSSION

The target monomer (**2**, **Scheme 1**), was synthesised from the previously reported[4] fluorenylmethoxycarbonyl (Fmoc)-protected amino alcohol (**1**) by oxidation with *bis*-acetoxyiodobenzene (BAIB) and 2,2,6,6-tetramethypiperidinyloxy (TEMPO) in aqueous acetonitrile. Although the oxidation proceeds in relatively low yield (46%) the method is advantageous in that the Fmoc-protected amino acid precipitates out of reaction mixture and is obtained pure after filtration and washing with diethylether.



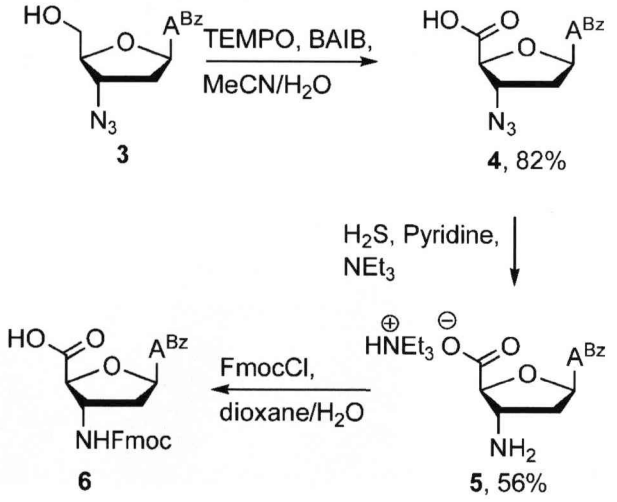
Scheme 1. Synthesis of oligomers derived from the thymidine beta-amino acid. T = thymin-1-yl; SPPS = solid-phase peptide synthesis.

Choice of resin for Fmoc solid-phase peptide synthesis (SPPS) was influenced by the sensitivity of the glycosidic bond to acid. With this in mind, the hyper acid-labile Sieber amide resin was chosen which on cleavage, produces peptides with a primary amide at the carboxylate terminus. The peptide assembly used HBTU in *N,N*-dimethylacetamide as the coupling agent and followed standard Fmoc protocols. Using this procedure dimers, tetramers and octamers, both with and without the Fmoc

group, were prepared (**Scheme 1**). In all cases the oligomers were sufficiently pure for analysis following cleavage from the support and washing with water.

Structural NMR studies were performed on both the Fmoc-protected tetramer and the deprotected octamer. Detailed NMR studies on the tetramer established a series of nOe conectivities which were utilized to build a model using Macro-Model. The model produced was consistent with an 8-membered ring between the backbone amide groups corresponding therefore to an "8-helix". This helical arrangement places the thymine bases on the outside of the helix where they could potentially H-bond to a complementary adenine base. With this in mind, we have therefore embarked on the synthesis of the corresponding adenine-derived oligomers in order to explore the possibility of helix association driven by base pairing.

The synthesis of the 2'-deoxyadenosine monomer (**6**) starts from the previously reported azidonucleoside[5] (**3, Scheme 2**) and utilizes benzoyl protection of the adenine base. In this case, the most efficient reduction of the azide was obtained using hydrogen sulfide. Current studies are underway on the solid-phase synthesis of peptides derived from adenine monomer (**6**).



Scheme 2, Synthesis of 2'-deoxyadenosine-derived beta-amino acid.

CONCLUSION

In conclusion, peptides derived from a thymidine beta-amino acid have been prepared and their conformation studied by NMR. Interestingly, the tetramer forms an 8-helix, which differs considerably from the 12-helix derived from the related *trans*-2-aminocyclopentanecarboxylic acid. Studies are currently underway on the synthesis of the complementary adenine peptides in order to investigate association between helices capable of nucleobase-pairing, such systems would offer a new motif in the search for innovative macromolecular architectures

Acknowledgements

The authors would like to thank the Daiwa Anglo-Japanese Foundation for financial support.

REFERENCES

1. Cheng, R.P., Gellman, S.H. and DeGrado, W.F. (2001) *Chem. Rev.*, **101**, 3219-3232.
2. Applequist, J., Bode, K.A., Appella, D.H., Christianson, L.A. and Gellman, S.H. (1998) *J. Am. Chem. Soc.*, **120**, 4891-4892.
3. A communication on part of this work has recently been published, Threlfall, R. Davies, A., Howarth, M.N. Fisher, J. and Cosstick R. (2008), *Chem. Commun.*, 585-587. Since our original publication, a report on the solution synthesis of the thymidine oligomers has appeared Chandrasekhar, S., Reddy, G.P.K., Kiran, M.U., Nagesh, C. and Jagadeesh, B. (2008), *Tet. Lett.*, **49**, 2969-2973.
4. Mignet, N. and Gryaznov, S.M. (1998) *Nucl. Acids Res.*, **26**, 431-438.
5. Pongracz, K, and Gryaznov, S.M. (1998) *Nucl. Acids Res.*, **26**, 1099-1106.

*Corresponding author. E-mail: rcosstic@liv.ac.uk